



LogView Pro 4 Beta 1  
Manual

# Contents

1	Introduction.....	4
1.1	What is LogView Pro.....	4
1.2	What's new in LogView Pro 4.....	5
1.3	Features.....	6
1.4	How it works.....	7
2	Setup.....	8
2.1	Installation.....	8
2.2	Preparing log files.....	8
2.3	Supported Log File Formats.....	9
2.4	Log Fields Specification.....	10
3	Search Expressions.....	11
3.1	Search box.....	11
3.2	Complex Expression Syntax.....	12
4	Configuration.....	13
4.1	Adding a new website.....	13
4.2	Main settings.....	14
4.3	Analyzer Settings.....	16
4.4	Counters.....	17
4.5	Parameters.....	19
4.6	Parameter extraction with Regular Expressions.....	20
4.7	Advanced Options.....	21
4.8	Import and Export Configuration.....	23
5	Working with Kainet LogView Pro.....	24
5.1	Session Overview.....	25
5.2	Traffic Statistics.....	27
5.3	Traffic Summary.....	28
5.4	Session Details.....	29
5.5	Session Pages.....	32
5.6	Session Page Details.....	33
5.7	Session Summary.....	34
5.8	WHOIS.....	35
5.9	RDAP.....	36
5.10	Trace Route.....	37
5.11	IP Address List.....	38
5.12	Useragent List.....	39
5.13	Country List.....	40
5.14	Referrer List.....	41
5.15	Bot Scores List.....	42

6	Filtering.....	43
6.1	Defining a Filter.....	43
6.2	Sub-Filter.....	48
6.3	Search for user names.....	49
6.4	Working with Filter Presets.....	50
6.5	Filter Combinations.....	51
7	Copy commands.....	52
8	Settings.....	53
8.1	Known Bots.....	53
8.2	Networks.....	54
8.3	WHOIS / RDAP settings.....	55
8.4	Other Settings.....	56
9	Strategies.....	58
10	Useful information.....	59
10.1	Logging basics.....	59
10.1.1	W3C Logfile fields / columns.....	59
10.1.2	Methods.....	61
10.1.3	Query Strings.....	62
10.1.4	Status Codes.....	63
10.1.5	Cookies.....	64
10.1.6	Sessions (Visits).....	65
10.1.7	DNS.....	65
10.1.8	WHOIS.....	65
11	Other related software.....	66
11.1	Kainet Editor.....	66
11.2	Kainet Aspi.....	67
11.3	CFLogfileConverter.....	68
11.4	LogFileDownloader.....	68

# 1 Introduction

## 1.1 What is LogView Pro

In short, Kainet LogViewPro is a Windows desktop software that visualizes the contents of web server logs and makes it easy to interactively explore and analyse the traffic on a web site.

The screenshot displays the Kainet LogView Pro interface. The top section shows a timeline from 19:45 to 00:00 with a grid of log entries. The bottom section shows a detailed table of request data for a specific IP address.

Time	Method	URL	Query	Statu	Referer	Bytes	Dur	Bot	Custom Fields
16.09.2024 16:11:48	GET	/entry-pages-assets/assets/js/a...		200	https://www.immoscout24.ch/...	140732	42	82	6730
16.09.2024 16:11:48	GET	/entry-pages-assets/assets/js/c...		200	https://www.immoscout24.ch/...	371278	80	82	6730
16.09.2024 16:11:48	GET	/de/wohnung/mieten		200	https://www.google.com/url?q...	38282	190	82	6730
16.09.2024 16:11:48	GET	/cdn-cgi/challenge-platform/s...		302		471	10	82	6730
16.09.2024 16:11:49	GET	/geo/locations-by-id	ids=geo-zipcode-5430	200	https://www.immoscout24.ch/	1590	90	99	6730
16.09.2024 16:11:49	POST	/search/listings		200	https://www.immoscout24.ch/	1174	99	96	6730, Log API traffic containir
16.09.2024 16:11:49	POST	/search/listings		200	https://www.immoscout24.ch/	1174	120	96	6730, Log API traffic containir
16.09.2024 16:11:49	GET	/cdn-cgi/challenge-platform/h...		200		4189	9	82	6730
16.09.2024 16:11:49	POST	/cdn-cgi/challenge-platform/h...		200		1073	37	82	6730
16.09.2024 16:11:49	OPTI...	/search/listings		204	https://www.immoscout24.ch/	1355	70	97	6730
16.09.2024 16:11:49	OPTI...	/search/listings		204	https://www.immoscout24.ch/	1355	70	97	6730
16.09.2024 16:11:50	POST	/cdn-cgi/rum		204	https://www.immoscout24.ch/...	380	4		6730
16.09.2024 16:11:55	GET	/geo/locations		200	https://www.immoscout24.ch/...	1075	00	00	6730

Summary statistics at the bottom: Swiss Web User\* 16'692 Sessions 19'447 Page Views 161'729 API Calls 1'008'294 Hits 0.75% Errors

In comparison to traditional, mostly cloud-based tools and solutions, this approach makes it possible to detect and identify low-key threats that otherwise would be overlooked.

One key advantage is the high responsiveness, which allows the user to perform much more analysis steps in much shorter time. This goes hand in hand with the idea of packing as much information as possible into meaningful graphical output.

## 1.2 What's new in LogView Pro 4

The first version of LogView Pro was developed in 2001, still using old technologies and languages. It still had a very basic functionality. With version two in 2004 the Software was rewritten under .NET 1.1 in C#. This was the basis for the following years.

For the present version four, the software was again completely revised, and about 95% of the code was completely rewritten. Especially the core libraries for traffic analysis now offer more flexibility and performance.

The software offers the following improvements:

- ◆ Performance improvements by a factor of up to three are achievable.
- ◆ Full support of IPv6 for all the functions
- ◆ Complex search expressions can be used for any text search
- ◆ Search for percentage of hits
- ◆ New table view with two display modes for session details replaces the old graphical view which fits better modern web applications
- ◆ Filter preset manager
- ◆ RDAP IP lookup in addition to the existing WHOIS queries
- ◆ IP and Useragent lists as tables
- ◆ Many small improvements in the UI that improve efficiency
- ◆ The libraries can be used programmatically to automate analyses and reports
- ◆ Support of Custom Fields

Overall, the new version keeps the features that proved useful in practice, and offers improvements that make log file analysis even more efficient and fun.

## 1.3 Features

Kainet LogView Pro allows visual inspection and analysis of web server logfiles. It is optimized for the following use cases:

- ◆ Identification of malicious traffic such as hack attempts or bad bots
- ◆ Criminal investigations
- ◆ Investigation of technical issues (errors, performance problems etc.)
- ◆ Plausibility check of numbers delivered by other systems

It is designed to work in an environment as follows:

- ◆ Log files have to be physically present locally for performance reasons
- ◆ Separate tools for download and conversion (for example CloudFlare log push) or merging (when load balancing is used) are available
- ◆ The W3C format (such as written by IIS or other servers) is directly supported
- ◆ Other formats will be supported in the future
- ◆ Tested with logs of over 15 GB and more than 30M entries
- ◆ Optimized for analysis of single days. Longer and shorter periods are possible

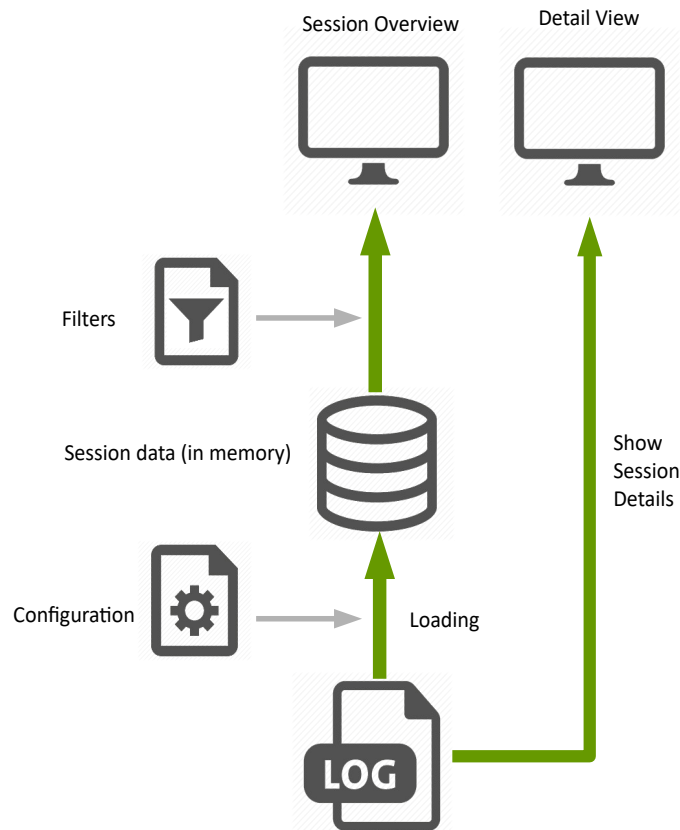
Key features:

- ◆ Session identification based on IP address, User-Agent and (optionally) Session Cookies
- ◆ Definition of counters for specific events per session
- ◆ Extraction of parameters per session
- ◆ Graphical overview over the entire or filtered traffic
- ◆ Different detail views for sessions
- ◆ Filtering of sessions with various standard and custom filter criteria
- ◆ Definition of text filters based on Regex, simple string search with wildcards or complex expressions with boolean operators
- ◆ Combination of filters, include and exclude

The feature set was developed over years based on practical experience.

## 1.4 How it works

Web server logfiles can get really huge and are usually difficult to handle. It's even more difficult to get useful information out of the millions of log entries. LogView Pro offers a solution to this problem. It makes it easy to open even huge logfiles and offers an interactive and quick visual approach to the desired information.



This is how it works:

- ◆ Before a log file can be loaded and analysed, a **configuration has to be defined** per website. This configuration contains information about where the logfiles are stored, which format, how to identify user sessions and what defines a page or API call. In this configuration we also can define counters for specific events, as well as definitions for parameter extraction
- ◆ Using this configuration, the **session data is loaded into memory**
- ◆ Using **filters** we can select the sessions we want to display
- ◆ To display **detailed information about each session**, LogView Pro retrieves the data directly from the original logfile. That's why we need local files.

This way we get access to all the information inside the logfile very intuitively and quickly.

## 2 Setup

### 2.1 Installation

To install LogView Pro and get advantage of the full functionality, just follow these steps:

- ◆ **Run the Setup** program. The software is now already functional, but the country information for each IP and session will still be missing.
- ◆ **Download an IP-to-country database.** Such can be downloaded for free here: <https://db-ip.com/db/download/country> (CSV version) We recommend this version, as the quality is sufficient, it is free, it is still relatively small and ensures good performance.
- ◆ **Extract** the file to any location on your local hard drive.
- ◆ Start LogView Pro and go to **File / Import Country database.** Choose the extracted CSV file and import it

Now LogView Pro is ready to go.

### 2.2 Preparing log files

For best performance, it is recommended to copy the web server logfiles to the local hard drive:

- ◆ Create a **separate folder** for each website
- ◆ **Download** the log files to these folders

For frequent analysis, this process can be automated.

We can provide separate command line tools to convert logfiles, merge them from a load balanced environment, or download and convert log data from platforms such as CloudFlare.



## 2.3 Supported Log File Formats

Currently the following formats are supported:

### W3C Extended

This is the format that is by default written by Microsoft IIS, but is also used by other web servers and logging modules.

- ◆ One line per entry (continuation of a line on a new line after an interruption according to standards is supported)
- ◆ Fields separated by white space
- ◆ The fields themselves must not contain white spaces. Those should be replaced by +
- ◆ The header contains all the information about the fields written.
- ◆ Field names according to W3C standards
- ◆ Custom fields for IP Address (or Forwarded IP address), Session ID, Country and BotScore are supported, the name of the field can be chosen freely
- ◆ For no value fields (null), a - should be written

The + character in certain fields (user agent, username, forwarded IP and Cookies) will be automatically replaced back by a white space. In theory, some information could get lost this way, but in practice this is almost never an issue.

We also recommend this format when writing conversion programs to use LogViewPro for analysis, as it is very fast to read.

### W3C Extended (compatible)

This format supports the full W3C standard. The differences to the previously described default format are:

- ◆ Strings can be put into quotation marks (quotes inside a string have to be doubled)
- ◆ Inside strings with quotes, white spaces can be used and don't have to be replaced by +

For details see:

<https://www.w3.org/TR/WD-logfile>

While reading this format, + characters are not replaced by white spaces.

Other formats will be supported in the future.

## 2.4 Log Fields Specification

The following fields are supported:

#Field W3C	Cloudflare Log Push	Content	Example
<b>date</b>	EdgeStartTimestamp (date)	Date of request (required if rollover is not at midnight)	2023-06-18
<b>time</b>	EdgeStartTimestamp (time)	Time of request	23:59:42
<b>c-ip</b>	ClientIP	Client IP ( <b>reverse proxy: use X-Clientip or X-Forwarded-For</b> )	10.20.0.1
c-port	ClientSrcPort	Client source port, can be used for criminal investigations	19094
cs-username	[custom field]	Username, if available	hans.muster@gmx.ch
<b>cs-method</b>	ClientRequestMethod	Request Method	GET
<b>cs-uri</b>	ClientRequestURI	Complete requested URI	/sample.php?id=1&name=hans
<b>cs-uri-stem</b>	ClientRequestURI (before ?)	Stem of requested URI	/sample.php
<b>cs-uri-query</b>	ClientRequestURI (after ?)	Query string of requested URI	id=1&name=hans
<b>sc-status</b>	EdgeResponseStatus	Response status	200
<b>sc-bytes</b>	EdgeResponseBytes	Size of response in bytes	48359
time-taken	OriginResponseDurationMs	Time taken by request in ms (useful for performance analysis)	474
<b>cs(User-Agent)</b>	ClientRequestUserAgent	Useragent Header	Mozilla/5.0 (Windows+NT..
cs(Cookie)	Cookie	If all cookies are logged, the session id can be extracted from the cookies	
cs(Referer)	ClientRequestReferer	Referrer Header	https://www.xy.com/home
cs-host cs(Host)	ClientRequestHost	Host Header (important if logfile is for different hosts)	www.xy.com
<b>[Session ID]</b>	Cookie: [custom field]		
[Bot Score]	BotScore	Bot Score from WAF	23
[Country]	ClientCountry	2-digit country code. If available, no country db is needed	CH
[Forwarded IP]		Similar X-Forwarded-For header (if behind reverse proxy, similar	178.197.208.199

Notes: Fields in bold are required or highly recommended. Special cases:

**IP Address:** By default, the field c-ip is used. But in case the server is behind a reverse proxy, the client IP can be taken from a custom field (for example X-Forwarded-For), which can also contain multiple IPs if behind multiple reverse proxies.

**Session ID:** The session ID can either be extracted from the cookies, or be taken from a custom field. The session ID helps with session separation, but LogView Pro also works without.

## 3 Search Expressions

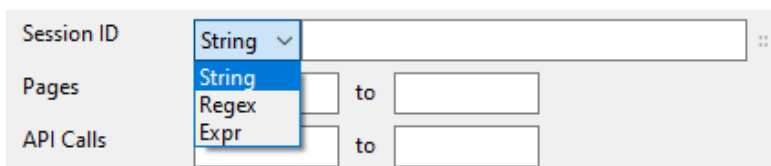
The concept of search expressions is important to understand, as it is used everywhere in the software where there's a search for a string.

There are **three ways** to search for a string:

- ◆ Simple string pattern
- ◆ Regular Expression
- ◆ Complex search expression

### 3.1 Search box

Wherever there is the option to use search expressions, there is a search box like this:



The options are:

String	<p>Simple string pattern that can contain the <b>wildcard * at the beginning or end of the string</b>. Wildcards in the middle are ignored and handled like normal characters. This limitation has its cause in performance considerations.</p> <p>Simple string patterns are easy to use and in many cases sufficient. For example to search for a certain user agent, just paste it without changes.</p>
Regex	<p>A <b>Regular Expression</b> is used to evaluate if a string matches the criteria. Regex offers the highest flexibility, but is also sometimes not easy to write.</p> <p>We recommend to test the regular expressions with an external tool first (for example Expresso)</p>
Expr	<p><b>Complex expressions</b> using boolean operators can be defined (see next chapter).</p>

## 3.2 Complex Expression Syntax

The following operators are supported:

&	AND
	OR
!	NOT
( )	Brackets with any number of levels
" "	Quotation marks have to be used if a pattern contains any of the operators above, or to search for white spaces
~"	The tilde sign preceding a quotation mark is used to specify a quotation mark inside quotation marks
@	At the beginning defines a regular expression within the complex expression

Example:

```
@"^Mozilla(?=.*Chrome)"&!("*49.0*"|*50.0*"|*Safari/537.36)
```

Will produce the following expression:

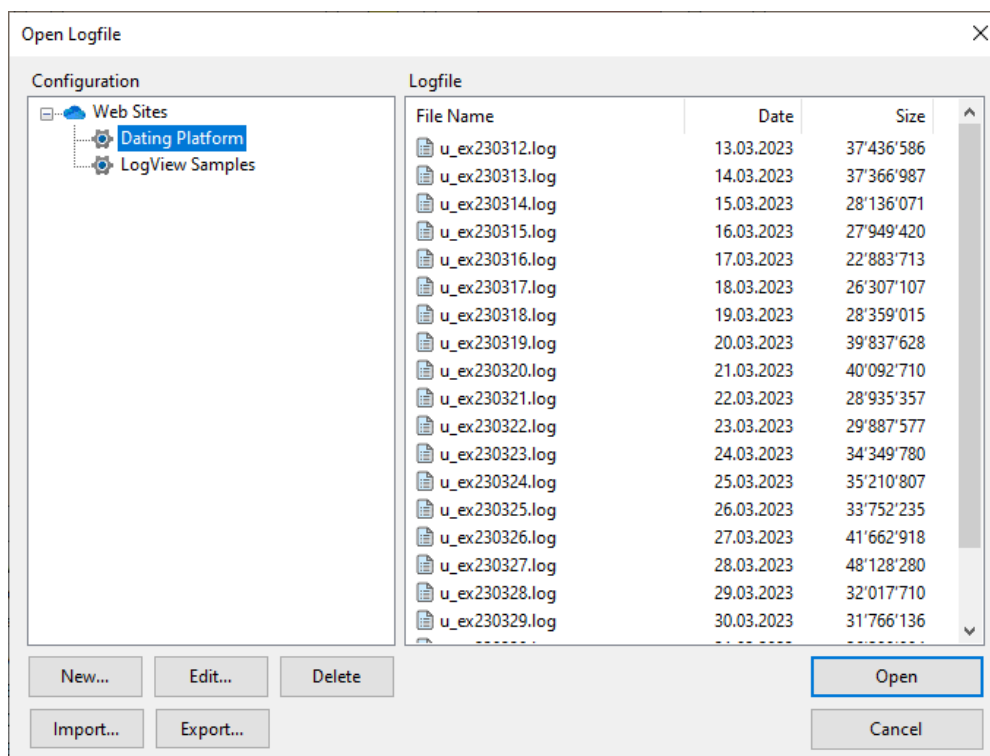
```
{
  "And": [
    {
      "RegularExpression": "^Mozilla(?=.*Chrome)"
    },
    {
      "Not": {
        "Or": [
          {
            "Contains": "49.0"
          },
          {
            "Contains": "50.0"
          },
          {
            "Pattern": "*Safari/537.36"
          }
        ]
      }
    }
  ]
}
```

To search for NULL, we can use the expression !\*

## 4 Configuration

### 4.1 Adding a new website

Before the first logfile can be opened and analysed, LogView Pro needs some information about the format, content and the web server configuration. This configuration is done in the window for opening log files which is accessed using **File / Open Logfile**:

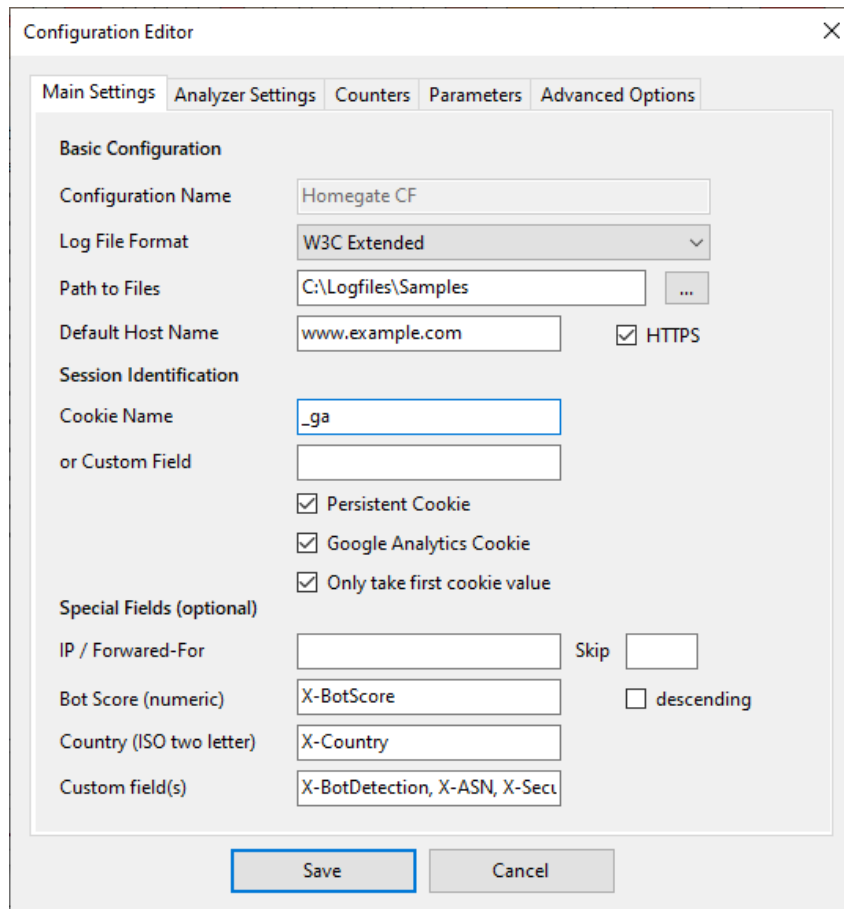


On the left side there is a list of websites that already have been configured. Initially, this list is empty. On the right side there is the list of all log files available configuration selected on the left.

To create a new configuration, just press **New...**

## 4.2 Main settings

The basic settings contain information required to load logfiles correctly:



The highlighted settings are required:

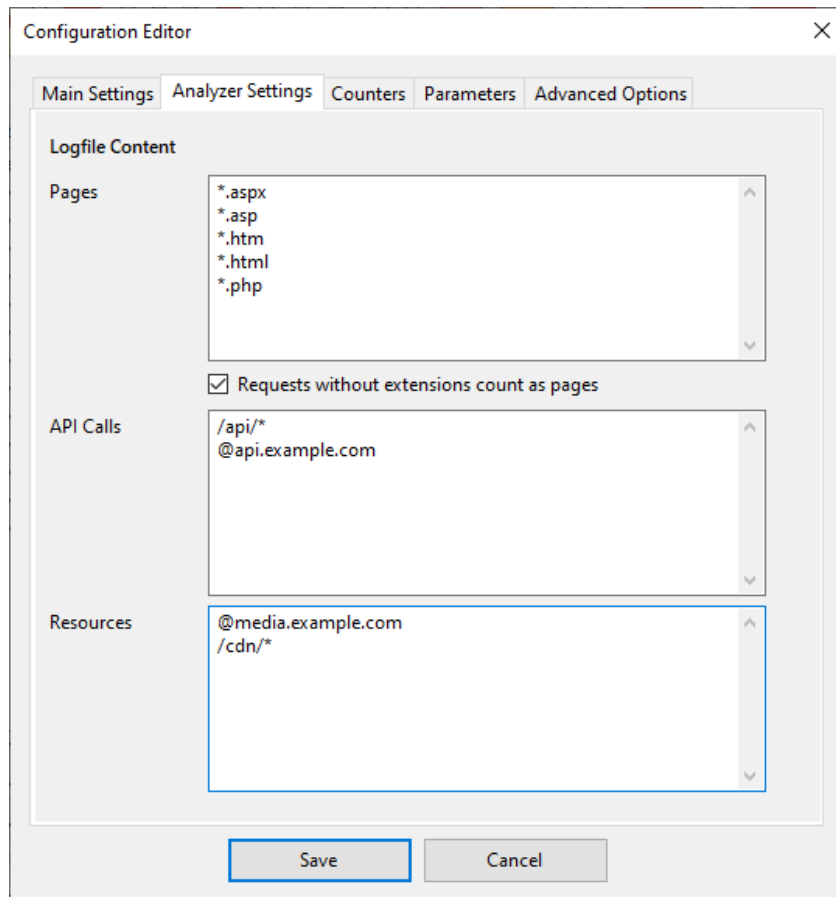
<b>Configuration Name</b>	Name of the configuration. Cannot be changed later.
<b>Logfile Format</b>	Web server log format. Formats currently supported are: <ul style="list-style-type: none"> <li>◆ W3C Standard (IIS)</li> <li>◆ other formats will be supported soon</li> </ul>
<b>Path to files</b>	Physical path to logfiles.
<b>Default Host Name</b>	The host name the website is running under
<b>Cookie Name</b>	Name of the session cookie to use to identify session, if cookies are logged
<b>or Custom Field</b>	Alternatively set a custom field in the log file where the session identification is stored (for example X-Cookie)
<b>Persistent Cookie</b>	Check if the session identifier is persistent (does not change over different session). The correct setting improves the accuracy of the session identification

Google Analytics Cookie	A Google Analytics cookie is used to identify sessions or users
IP / Forwarded for	If the server resides behind a reverse proxy, the information about the original IP is stored in a separate field, stemming from a separate HTTP header (for example X-Forwarded-For)
Skip (numeric)	If the server resides behind multiple reverse proxies, one or more entries can be skipped to get the correct originating IP
Bot Score (numeric)	Some firewalls or WAFs calculate a bot score, the likelihood that a request is originating from a bot. The corresponding log field can be specified here.
Descending	With default (off), zero indicates the highest likelihood that it's a bot. When checked, its the lowest
Country	If the logfiles contain country information in a custom field (for example X-Country, it can be specified here. This will override the built-in country database.
Custom Fields	<p>If the logfile contains other fields of interest, those can be included here. It is possible to define one or multiple fields.</p> <p><b>Important:</b> The values from all the custom fields are collected into a single list per session, for which the sessions can be filtered. Multiple different values within the same sessions are all added to the list, while identical values are stored just once.</p> <p>So it is important not to add fields that have different values for each request. Otherwise performance issues could occur.</p>

After entering the basic settings, the first log file could already be loaded. But we recommend to also do the next step before.

## 4.3 Analyzer Settings

These settings are important for the software to display and analyse traffic correctly. LogView Pro needs to know how to **recognize what is a page, an IP call**, and what is just a **resource**, like a picture or any static asset.



The wildcard character `*` can be used at the beginning or end of a pattern. It is also possible to use a host name instead by using `@` at the beginning. In this case no wildcards are allowed.

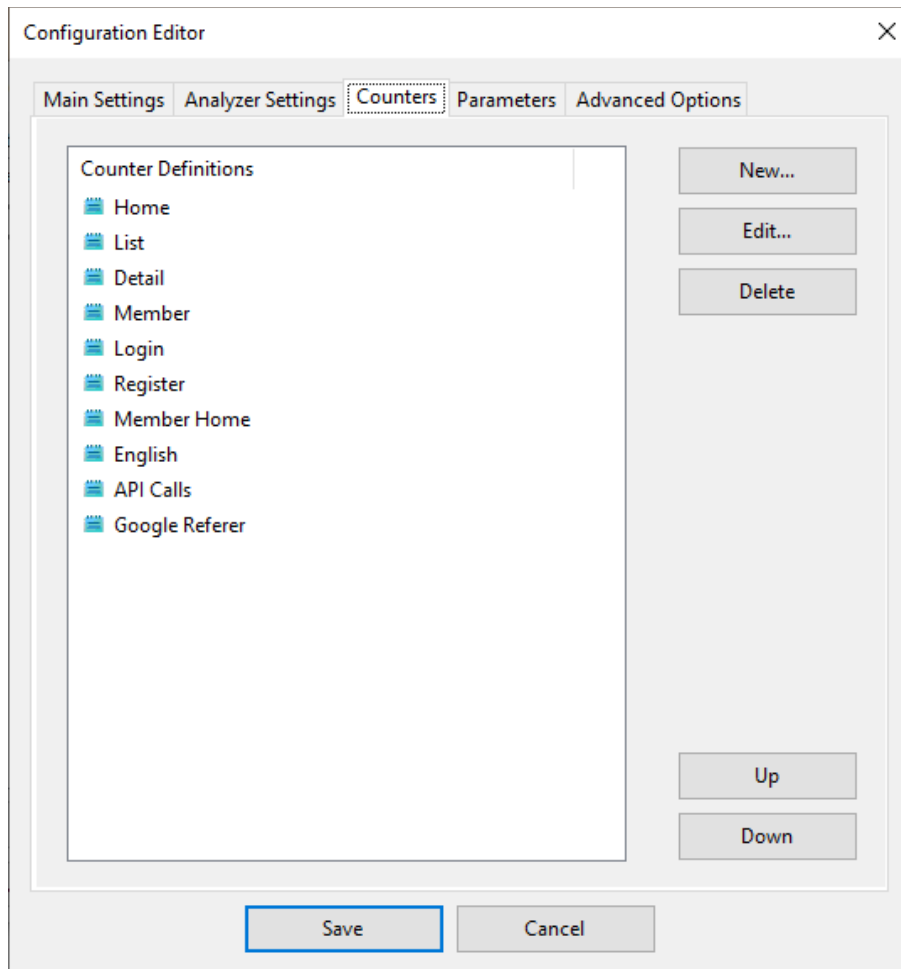
Pages	All patterns that define a full web page (such as a HTML page or e server side rendered page). The patterns can just contain a file ending, but also whole paths
Requests without extension count as pages	Nowadays the practice of rewriting URLs is common. So usually a web page is called without a suffix such as <code>.html</code> . Checking this box defines that every request without a suffix are considered to be web pages. Exceptions can be defined below.
API Calls	Patterns (suffixes or paths) that identify API calls
Resources	Patterns (suffixes or paths) that identify resources (neither API calls nor page requests)



## 4.4 Counters

Counters will **count occurrences** of requests matching the criteria **for each session**. So later sessions can be filtered by the number of occurrences:

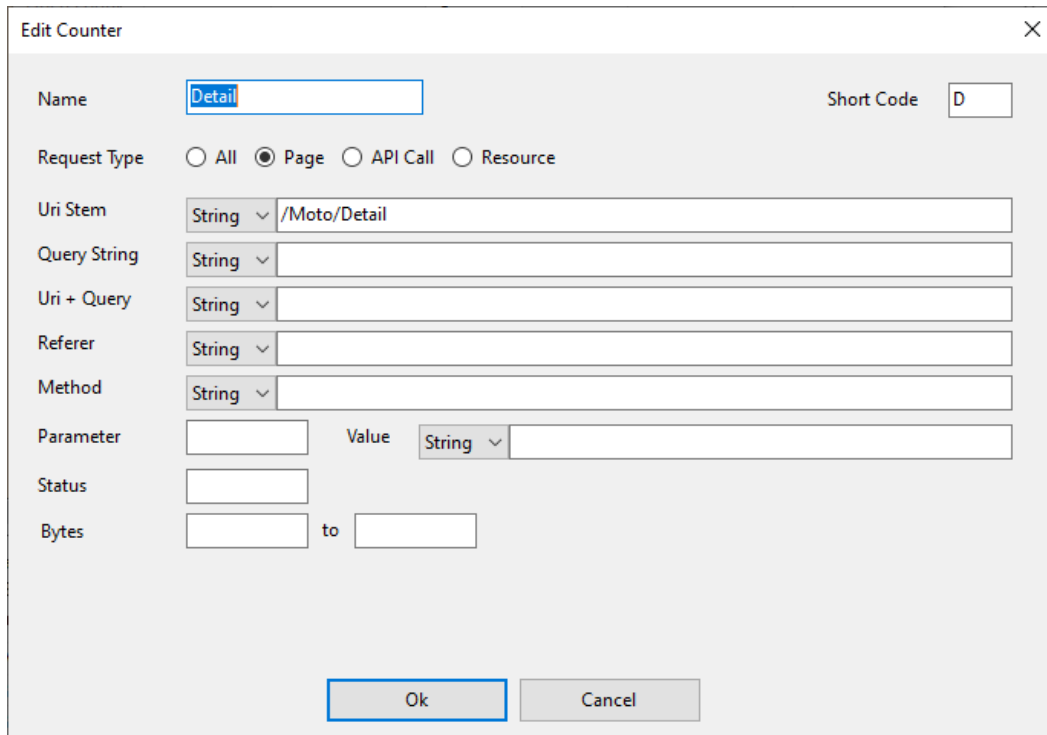
For each request it is checked whether the requests matches the criteria in the filter definition. If the criteria match, the request is counted.



To define a new counter, just press **New...**

On this tab, counter definitions can be edited, deleted or the order can be changed.

A filter definition looks as follows:



The criteria are combined with **AND**. Fields that are left empty will not be considered as criteria.

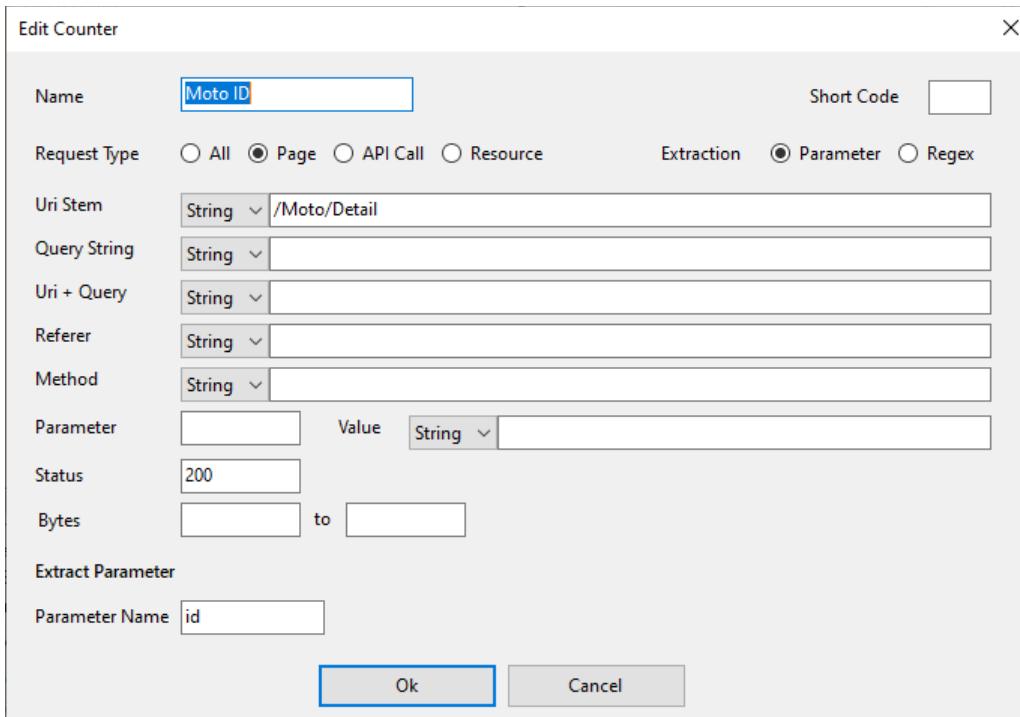
<b>Name</b>	Unique name of a counter definition. The name can be changed later.
Short Code	A short code can be defined that will improve readability of the session detail view. Ideally a short code is <ul style="list-style-type: none"> <li>◆ Only one character long</li> <li>◆ Uppercase for important counters</li> </ul> In the session detail view rows can be <b>highlighted</b> based on the short codes, lower- or uppercase.
Request Type	Limits the counter to certain request types (defined in Analyzer Settings). This helps to improve performance, but can also be helpful in certain cases
URI Stem	Stem part of the URL (without the query string). A search expression can be used.
Query String	The query string part of the URL (after ?)
URI + Query	The search expression covers both URI stem and query string
Referer (sic!)	Referer header of the request
Method	Typically GET, POST, sometimes PUT, DELETE etc.
Parameter	The request has to contain a certain parameter: The parameter name is just a string, while the value has to match a search expression

Status	HTTP Status code (numeric), for example 404
Bytes from / to	Range for the number of response bytes returned by the server.

## 4.5 Parameters

It is possible to extract values from a request (typically from the query string, but also from the stem part). After that, sessions can be searched for this value. Questions like this can be answered: “Who has looked at a certain object?”.

If the value is part of the query string parameters, the standard definition can be used:



The criteria are the same as for the counter(see previous chapter). Then the query string parameter will only be extracted if all the criteria match. All the fields can be left empty.

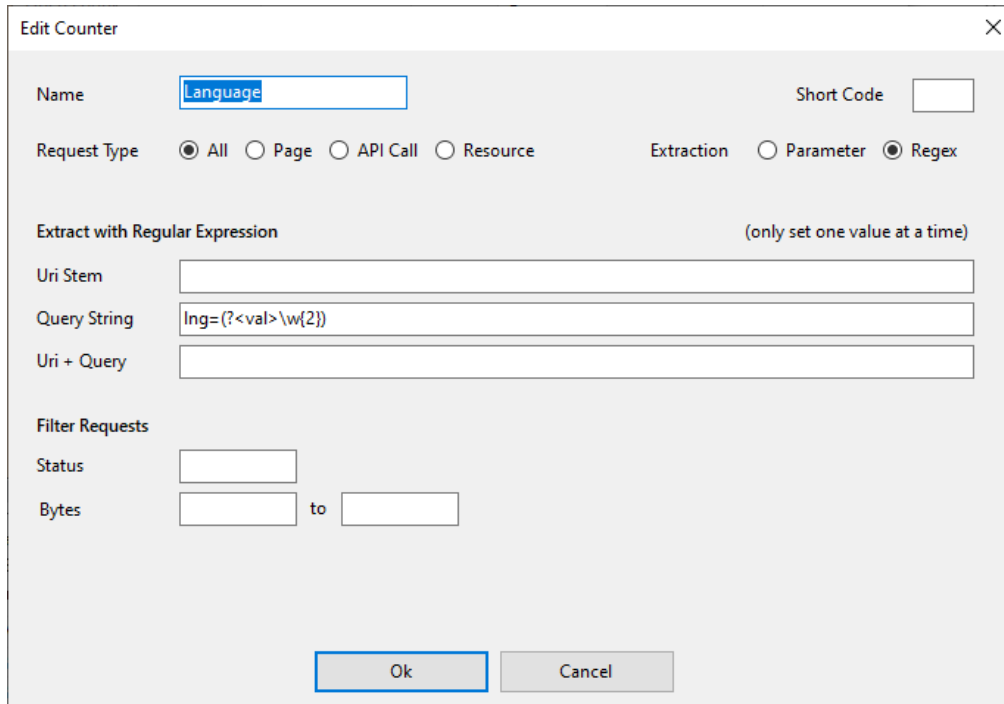
But additionally we have to define the name of the query string parameter we want to extract.

A short code can be defined, but is currently not used.

## 4.6 Parameter extraction with Regular Expressions

Additionally we have the possibility to use Regular Expressions to extract a value. Very often an ID or value is part of the URI stem, so Regex is very useful to extract any part of any URI.

We can switch to Regex extraction by choosing the **method on the top right**:



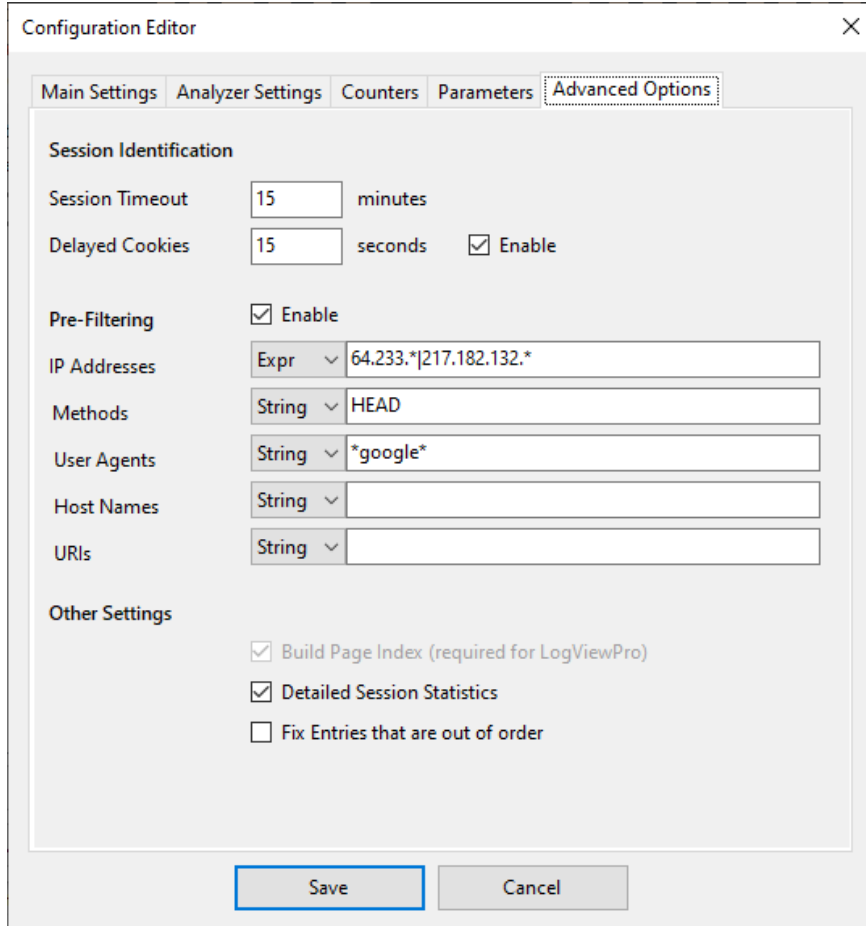
A regular expression can be defined for the URI stem, the query string, or the entire URL including the query string. But **only one of the three options can be used at the same time**.

A Regular Expression must contain a **capture** like **(?<name> ... )** where the name and the number of captures is irrelevant. **Just the first capture found is used**.

Additionally the requests can be **filtered** by status and / or number of response bytes.

## 4.7 Advanced Options

Some advanced options can be used to **fine-tune** the pre-analysis to suit best the specific characteristics of a website:



Configuration Editor

Main Settings | Analyzer Settings | Counters | Parameters | **Advanced Options**

**Session Identification**

Session Timeout: 15 minutes

Delayed Cookies: 15 seconds  Enable

**Pre-Filtering**  Enable

IP Addresses: Expr

Methods: String

User Agents: String

Host Names: String

URIs: String

**Other Settings**

Build Page Index (required for LogViewPro)

Detailed Session Statistics

Fix Entries that are out of order

Save Cancel

Session Timeout	The time interval after which a session is considered as terminated if no more requests are coming in. The default of 15 minutes suits best most practical use cases. In general, a longer timeout will result in a smaller number of overall sessions
Delayed Cookies	Usually a client receives an identifying cookie with one of the first requests. This also means that the first requests come in without the cookie. Additionally it is possible that during a certain period, some requests have the cookie, others again not. LogView Pro has a mechanism implemented to assign the incoming requests to the correct session anyway. We highly recommend to leave this feature enabled. The default setting of 15 seconds is usually the best compromise. But if a web application takes very long to return the cookie reliably, a higher value can be used.

---

Pre-Filtering	<p>In some cases it is not necessary to analyse the entire traffic that is contained in the log file. So traffic matching certain criteria can be excluded from the analysis completely.</p> <ul style="list-style-type: none"> <li>◆ IP Addresses (for example internal networks)</li> <li>◆ Methods (for example HEAD, OPTIONS)</li> <li>◆ User Agents</li> <li>◆ Host Names</li> <li>◆ URI pathes</li> </ul> <p>This pre-filtering has to be done using search expressions, as it happens very early, when all fields are only available as strings.</p> <p>Remember that everything that was filtered out here will be completely invisible.</p>
Build Page Index	<p>LogView Pro uses an internal index for request positions in the log file source. This cannot be disabled.</p>
Detailed Session Statistics	<p>To display detailed statistics on minute basis, traffic statistics data has to be saved for every session. This consumes memory and increases the load time.</p> <p>But we recommend to leave this option enabled, as long as there are no serious performance issues</p>
Fix Entries that are out of order	<p>LogView Pro assumes that all logfile entries are correctly sorted. A tolerance of +/- 5 seconds is accepted. If entries are out of order more than this, they are ignored and an error message is logged.</p> <p>With this option, entries can be corrected automatically, then the time is overwritten with a value that matches the order.</p> <p>This option should only be used if the number of entries out of order is low. Otherwise the logs should be pre-processed with an external tool. (For example the CFLogfileConverter is doing exactly this).</p>

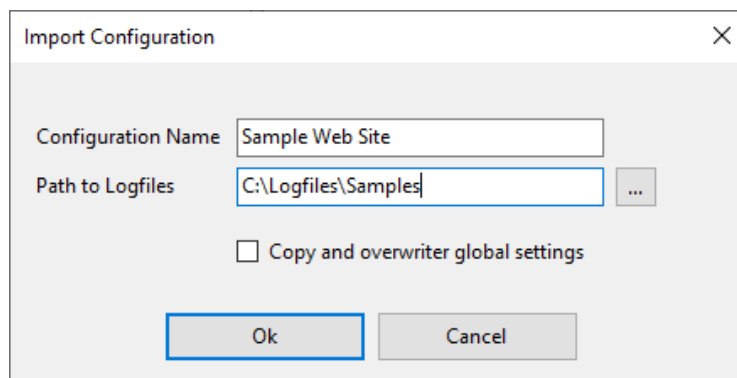
---

## 4.8 Import and Export Configuration

By default configurations are stored in the folder **C:\ProgramData\LogViewPro4**. Configuration files and the corresponding filter files can be copied manually from there to other systems.

To make the process safer and easier there is the option to export a complete configuration, including filters and global settings into one file. To do so, press **Export...** in the **Open Logfile** window.

To import a configuration on another system, there is the **Import..** function. The following window will be displayed:



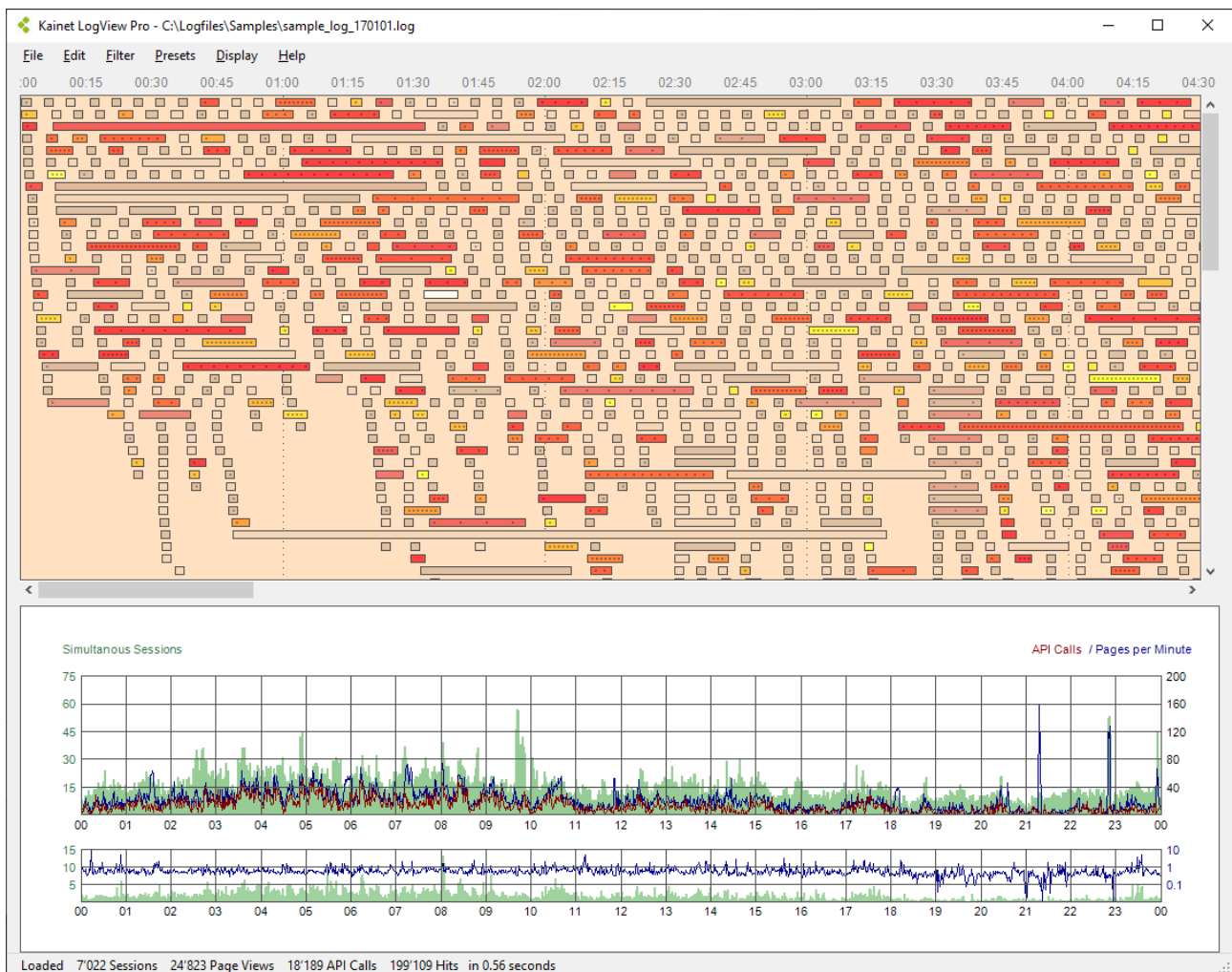
Here it is possible to **change the name of the configuration** and the **folder for the log files** while keeping all the settings and filters.

Optionally the global settings can also be imported. **But be careful:** This will overwrite all the global settings with the ones contained in the export file.

This function can also be used to duplicate configurations, for example for a similar website, or for special analyses.

## 5 Working with Kainet LogView Pro

After a logfile is loaded and pre-analysed, the traffic is visualized in the main window:



The main window is split into two parts:

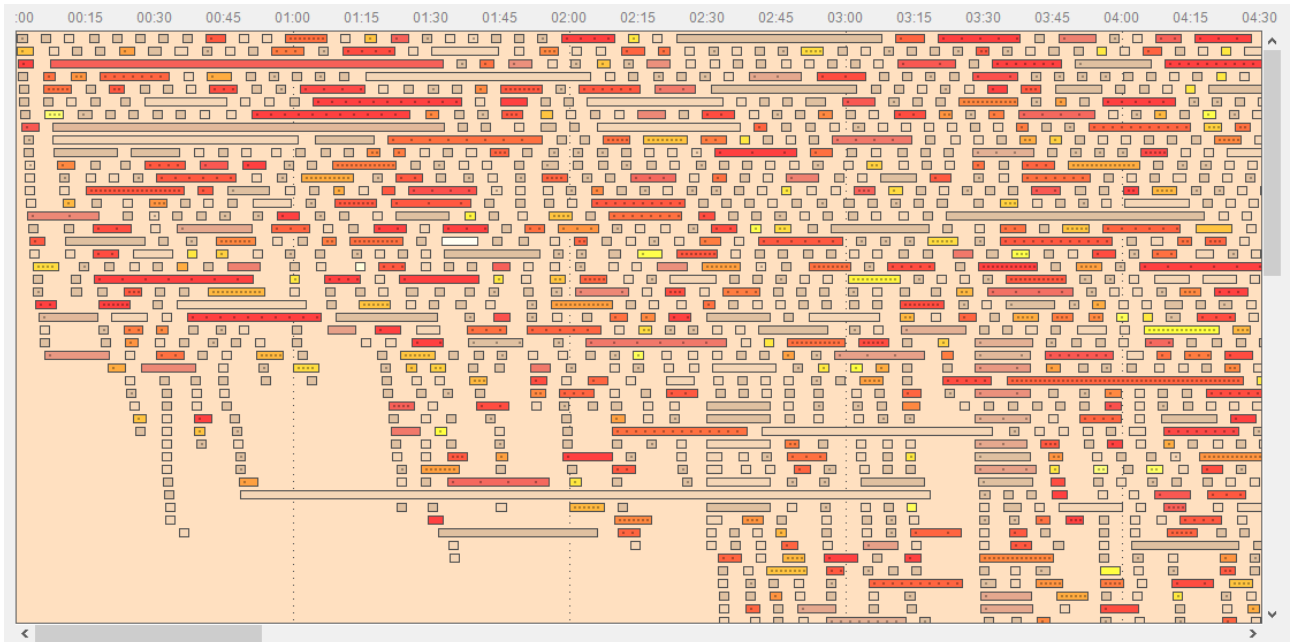
- ◆ Upper part: The session overview where all the loaded sessions are displayed as coloured bars in a timeline.
- ◆ Lower part: Different views on the entire traffic or individual sessions can be selected via the **Display Menu** or using hotkeys.

Both displays always show the currently active (filtered) sessions, or, for the lower part, the currently selected session, depending on the view mode.



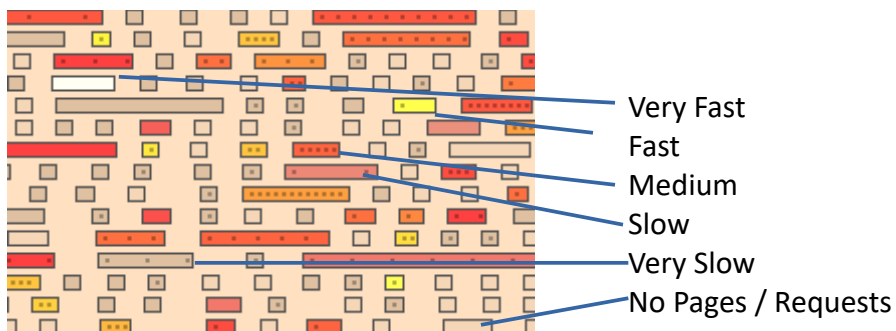
## 5.1 Session Overview

All the currently active sessions are rendered as clickable bars. First after loading, all the sessions in the log file are visible. Because these can be really a lot, it is useful to define a filter, so only the sessions of interest are displayed (see chapter about filtering).



The **colour** depends on the click speed inside the session, that means the number of pages per minute. Sessions with a very high rate are displayed in white, the slowest in grey. In general: The "hotter" the colour, the more requests.

Example:



There are **three different colour modes** which can be switched via the menu, or using hotkeys:

<b>Pages</b> (Ctrl + F1)	The <b>speed of the page requests</b> determines the colour, while the spacing of the dots is determined by the number of API calls per time
<b>API Calls</b> (Ctrl + F2)	Just the other way round: <b>API calls per minute for the colour</b> , Pages for the dots
<b>Hits</b> (Ctrl + F3)	The number of all <b>hits per minute</b> determine the colour, while Pages + API calls per time define the intervals between the dots.

The Session Overview can be navigated using the mouse wheel:

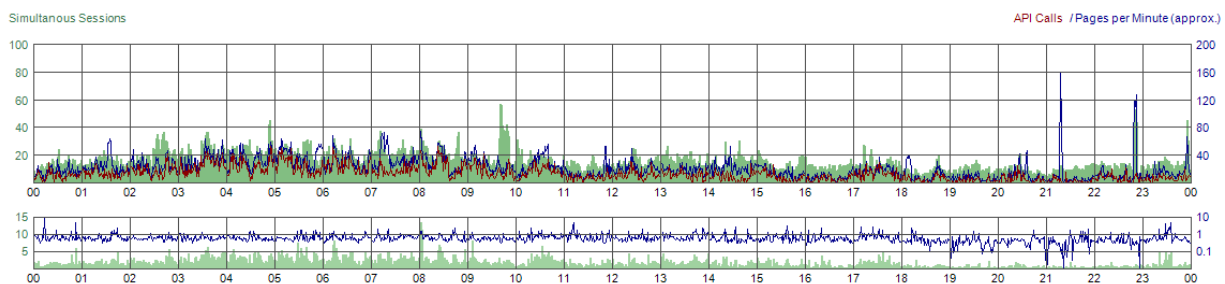
- ◆ Scroll wheel: Scrolls up and down
- ◆ Scroll wheel + Shift: Scrolls left and right

## 5.2 Traffic Statistics

The graphical statistics display shows an overview over the traffic for the entire length of the loaded log file. It always shows the numbers **for the current active filter**, that means for the currently visible sessions.

This display is a useful tool to find anomalies, for example bot traffic. By narrowing down the traffic using filters or filter combinations, peaks in any of the lines will get more distinct.

It works best if the log file contains around one day, but it hasn't to be from midnight to midnight.



Upper graph:

- ◆ The green filled area shows the **number of simultaneous sessions**
- ◆ The blue line indicates the number of **page requests per minute**
- ◆ The red line indicates the number of **API calls per minute**

Lower graph:

- ◆ The green filled area shows the number of **request** running on the server **simultaneously**. This is an approximation.
- ◆ The blue line shows the **average time used per request** in seconds. The scale is fixed but logarithmic, as usually the execution time lies between 0.01 and 10 seconds.

Note: Cloudflare measures zero ms for most requests, except if they last very long, so the lower graph might be not too useful.

The scale is adjusted dynamically.

**Copy (Ctrl + C)** copies the entire bitmap for use in a document.

The traffic statistics are displayed on start by default or can be activated via the **display menu** or by **pressing F1**

## 5.3 Traffic Summary

The summary can be displayed via the **display menu** or by pressing **F4**.

General Counters	Total	per Session	
-----			
Hits	199'109	28.4	
API Calls	18'189	2.6	9.1%
Pages	24'823	3.5	12.5%
Client Errors	1'484	0.2	0.7%
Server Errors	1	0.0	0.0%
Sessions	7'022		
IP Addresses	4'199		
Custom Counters	Hits	Sessions	
-----			
Home	6'913	2'141	3.5%
List	10'644	1'504	5.3%
Detail	9'532	2'913	4.8%
Member	890	149	0.4%
Login	245	121	0.1%
Register	41	26	0.0%
Moto ID (20 of 2473)	Hits	Sessions	
-----			
17000	517	475	
16999	39	34	
17016	34	33	
...			
Language	Hits	Sessions	
-----			
th	21'371	3'338	
en	2'674	802	

This display mode summarizes the numbers for the **currently filtered (visible)** sessions.

It contains:

- ◆ General counters: Hits, API calls, pages, client and server errors, sessions, different IP addresses
- ◆ Custom counters
- ◆ Top 20 of each captured parameters

The percentage is always relative to the number of hits.

## 5.4 Session Details

The Session Details view allows a look into all requests that are part of a session, that means to see everything a single user has done during his visit on the web site.

This view is **automatically selected** when a session in the traffic view is clicked and a summarized view was active (traffic view or traffic summary).

Switching to it manually can be done via the **display menu** or pressing **F5**.

Time	Method	URL	Query	Status	Referer	Bytes	Dur	Bot	Session ID
01.01.2017 03:34:48	H GET	/	gclid=CP2InlvPtNMCFVgfaAodt-cDGw	301	https://www.google.co.th/		384		GA1.2.177214
01.01.2017 03:34:51	H GET	/	gclid=CP2InlvPtNMCFVgfaAodt-cDGw	200	https://www.google.co.th/		634		GA1.2.177214
01.01.2017 03:34:52	GET	/api/Counter	ids=16979,16983,16991,17011,16962,1698...	200	https://www.kainetmotos.com...		308		GA1.2.177214
01.01.2017 03:34:52	GET	/Photos/Small/0064/671.jpg		200	https://www.kainetmotos.com...		366		GA1.2.177214
01.01.2017 03:34:52	GET	/Photos/Small/0064/821.jpg		200	https://www.kainetmotos.com...		323		GA1.2.177214
01.01.2017 03:34:52	GET	/Photos/Small/0064/726.jpg		200	https://www.kainetmotos.com...		718		GA1.2.177214
01.01.2017 03:34:53	GET	/Photos/Small/0064/593.jpg		200	https://www.kainetmotos.com...		663		GA1.2.177214
01.01.2017 03:34:53	GET	/Photos/Small/0064/703.jpg		200	https://www.kainetmotos.com...		625		GA1.2.177214
01.01.2017 03:35:08	H GET	/		200	https://www.kainetmotos.com...		1728		GA1.2.177214
01.01.2017 03:35:09	GET	/api/Counter	ids=16979,16983,16991,17011,16962,1698...	200	https://www.kainetmotos.com/		310		GA1.2.177214
01.01.2017 03:35:13	L GET	/Moto/List	lng=th&sort=latest&top=50	200	https://www.kainetmotos.com/		648		GA1.2.177214
01.01.2017 03:35:14	GET	/api/Counter	ids=17019,17018,17017,17016,17015,1701...	200	https://www.kainetmotos.com...		310		GA1.2.177214

The **header area** contains basic **information about the session**, such as IP Address, reverse DNS (if available), user agent, number of hits, API calls or page views and the duration.

If **Google Analytics cookies** are used to identify sessions, instead of the original cookie, just “GA” and the **creation date and time** of the cookie are displayed.

Each **row** in the table represents a **request** that is part of the session. These are the columns:

Time	Date and time of the request
Short Code	Short codes(s) as defined by the <b>counters</b> (see configuration)
Symbol	Represents type and response status of the request (see below)
Method	HTTP method
URL	Stem part of the URI
Query	Query string part of the URI
Status	HTTP response status
Referer	Referer(sic) header of the request
Bytes	Number of response bytes
Dur	Time taken by the request on the server side
Bot	Bot score (if available)







The content of the last column, called **Data Column**, can be switched via the menu:

Session ID	Value of the session cookie (if available)
Username	Username (if available)
Custom Fields	List of all the custom field values collected for that request

The symbols have the following meaning:

	Page request
	API Call
	Resource
	Redirect (status 3xx except 304)

The symbols can appear in different colours:

	Success, the action was successfully executed (status 2xx)
	No data returned (status 304 or methods OPTIONS, HEAD)
	Not found or gone (status 404 or 410)
	Other client error (status 4xx)
	Server error (status 5xx)
	Virtual page requests, reconstructed from referer (only in Session Pages View)

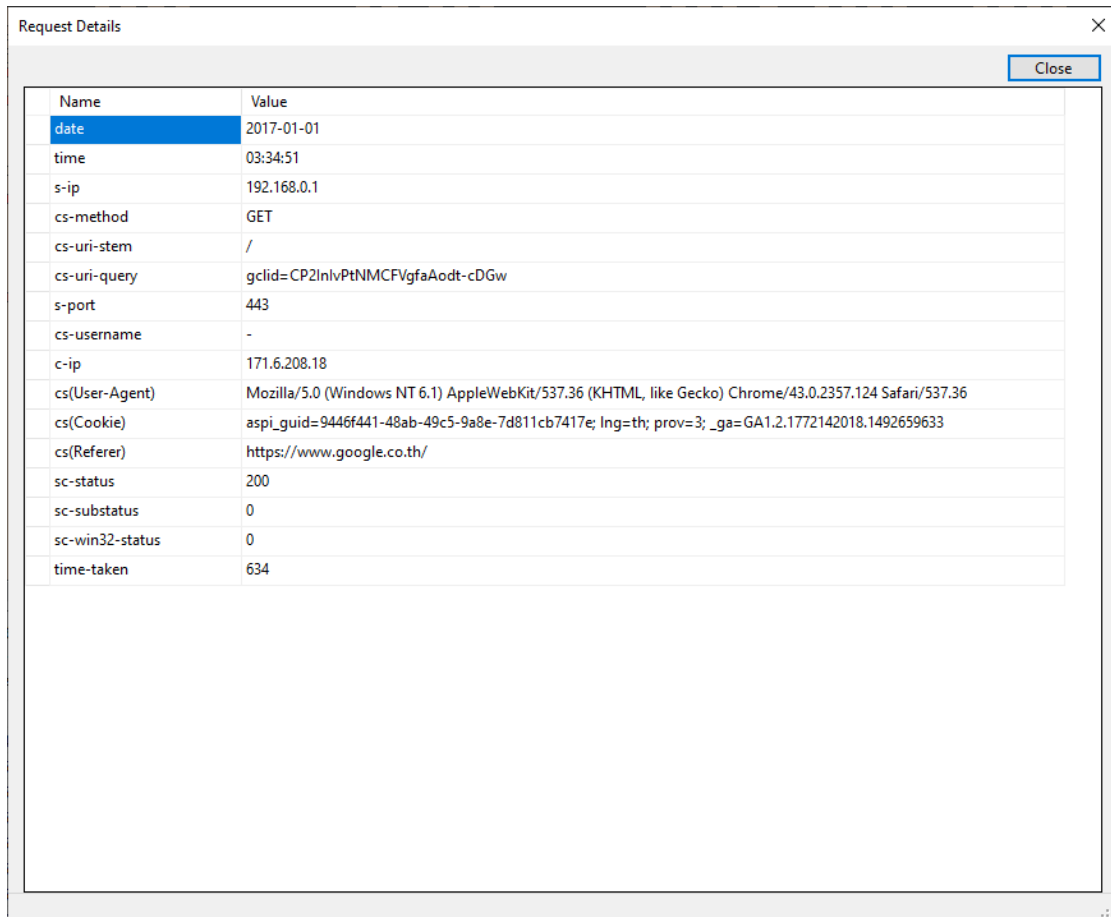
### Highlighting of rows

Rows can be automatically highlighted. Highlighting is enabled via the **Display menu**, or using the hotkeys **Ctrl+F5 to F6**

None	No rows are highlighted, white background
Short Code	All rows with any short code are highlighted
Short Code Uppercase	Only rows with a short code in upper case is highlighted

**Copy (Ctrl + C)** works with the table view too: All selected cells are copied to the clipboard.

All the details to a request can be displayed by **double clicking** any row:



The screenshot shows a window titled "Request Details" with a "Close" button in the top right corner. The window contains a table with two columns: "Name" and "Value". The "date" row is highlighted in blue.

Name	Value
date	2017-01-01
time	03:34:51
s-ip	192.168.0.1
cs-method	GET
cs-uri-stem	/
cs-uri-query	gclid=CP2InlvPtNMCFVgfaAodt-cDGw
s-port	443
cs-username	-
c-ip	171.6.208.18
cs(User-Agent)	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/43.0.2357.124 Safari/537.36
cs(Cookie)	aspi_guid=9446f441-48ab-49c5-9a8e-7d811cb7417e; lng=th; prov=3; _ga=GA1.2.1772142018.1492659633
cs(Referer)	https://www.google.co.th/
sc-status	200
sc-substatus	0
sc-win32-status	0
time-taken	634

In this window, the **original log fields** from the log file for that specific request are displayed, including fields that are not used by LogView Pro. This feature gives access to information that is available in the log file, but not part of any standard.

## 5.5 Session Pages

This view can be accessed via the **Display Menu** or by pressing **F6**. It basically only displays page requests, and allows too see what pages had been viewed during a user session. But this description is not fully accurate.

One line is displayed per:

- ◆ **Page request**
- ◆ Requests (API calls, resources, redirects) that are **not referred** by any page
- ◆ **“Virtual” page requests (in grey)**: If no real matching page request for any Referrer-Header in an API call or resource can be found, a “virtual” page request is created for every different referrer.

The “virtual” page requests are especially useful in **modern web applications**, where the URI of the page is set on the client side while client-side rendering the page. No real page is requested then, just API calls are executed.

The Referrer column is replaced by a **Subrequests**-column, where all the API calls and hits for a specific page are counted.

Please note that the assignment of the API calls and hits to each virtual page are **not completely accurate**, as usually the current URI is set by the web application **after** the first API calls are executed. Nevertheless the view is quite useful.

Time	Method	URL	Query	Statu	Subrequests	Bytes	Dur	Bot	Session ID
01.01.2017 03:43:25	H GET	/	gclid=Cj0KEQjwuOHHRDmvsHs8Pukyl...	200	1 API Calls 18 Hits	12509			GA1.2.19555
01.01.2017 03:43:27	GET	/Styles/Site.css			0 API Calls 1 Hits	588			
01.01.2017 03:48:08	H POST	/	gclid=Cj0KEQjwuOHHRDmvsHs8Pukyl...	302		315			GA1.2.19555
01.01.2017 03:48:08	L GET	/Moto/List	lng=th	200	1 API Calls 17 Hits	7193			GA1.2.19555
01.01.2017 03:51:40	D GET	/Moto/Detail	lng=th&id=17013&nr=9&tot=1612&q...	200	0 API Calls 12 Hits	8323			GA1.2.19555
01.01.2017 03:52:26	GET	/Moto/List	lng=th		0 API Calls 1 Hits	277			GA1.2.19555
01.01.2017 03:52:27	L GET	/Moto/List	lng=th&page=2	200	1 API Calls 11 Hits	3836			GA1.2.19555
01.01.2017 03:52:46	D GET	/Moto/Detail	lng=th&id=17008&nr=14&tot=1612&q...	200	0 API Calls 3 Hits	2090			GA1.2.19555
01.01.2017 03:52:58	L GET	/Moto/List	lng=th&page=3	200	1 API Calls 9 Hits	2920			GA1.2.19555
01.01.2017 03:53:12	L GET	/Moto/List	lng=th&page=4	200	1 API Calls 11 Hits	3405			GA1.2.19555
01.01.2017 03:53:19	L GET	/Moto/List	lng=th&page=5	200	1 API Calls 11 Hits	3467			GA1.2.19555
01.01.2017 03:53:33	L GET	/Moto/List	lng=th&page=6	200	1 API Calls 12 Hits	3960			GA1.2.19555
01.01.2017 03:54:00	L GET	/Moto/List	lng=th&page=7	200	1 API Calls 11 Hits	3774			GA1.2.19555

**Shortcut:** A easy way to **switch** between Details view and Pages view is by **double-clicking** on the empty space in the header area.



## 5.6 Session Page Details

The third detail view for sessions is a combination of both. One line is displayed for:

- ◆ **All requests** (pages, API calls and resources)
- ◆ **“Virtual” page requests** (in grey): If no real matching page request for any Referrer-Header in an API call or resource can be found, a “virtual” page request is created for every different referrer.

This view can be helpful in certain cases: All information is visible, the real requests, and the referred “virtual pages”.

Time	Method	URL	Query	Status	Referer	Bytes	Dur	Bot
26.07.2024 20:00:15	S	GET	/de/s/mo-330/mk-bmw	200	https://www.google.com/	60'947	0	98
26.07.2024 20:00:17	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:17	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:17	■	POST	/cdn-cgi/rum	204	https://www.autoscout24.ch/d...	380	0	
26.07.2024 20:00:17	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:17	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	703	0	98
26.07.2024 20:00:17	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	706	0	99
26.07.2024 20:00:17	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	99
26.07.2024 20:00:31			/de/s/mo-330/mk-bmw					
26.07.2024 20:00:31	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	703	0	98
26.07.2024 20:00:32	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	706	0	98
26.07.2024 20:00:34	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	10'231	0	98
26.07.2024 20:00:34			/de/s/advanced					
26.07.2024 20:00:34	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:34	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:34	d	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	701	0	98
26.07.2024 20:00:34	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	706	0	98
26.07.2024 20:00:39	■	POST	/cdn-cgi/rum	204	https://www.autoscout24.ch/d...	380	0	
26.07.2024 20:00:39			/de/s/advanced					
26.07.2024 20:00:39	s	GET	/listings-web/_next/data/861d7...	200	https://www.autoscout24.ch/d...	706	0	98

## 5.7 Session Summary

The session summary can be accessed via the **Display menu**, or by pressing **F8**. It shows a text based summary of all the information to the currently selected session.

Session Information		
-----		
IP Address	223.204.108.18 TH (Thailand)	
User Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:53.0) Firefox/53.0	
Session ID	GAX.x.144185910.1492740944	
Start Time	01.01.2017 02:15:41	
End Time	01.01.2017 02:16:50	
Duration	00:01:09	
First Request	/?gclid=Cj0KEQjwuOHHBRDmvsH	
Referer	https://www.google.co.th/	
Hits	65	
Pages	4	6.2%
API Calls	4	6.2%
Client Errors	0	0.0%
Server Errors	0	0.0%
Matching Filters		
-----		
Human Traffic		
from Google		
from Google paid		
MS and Facebook		
Counters		
-----		
Home	3	4.6%
List	3	4.6%
English	1	1.5%
API Calls	4	6.2%
Google Referer	2	3.1%
Language		
-----		
th	3	
en	1	

It contains the following information:

- ◆ **Basic information**, such as IP address, User agent, Session ID/Cookie (if available), time, duration, first request, referrer.
- ◆ **Standard counters**: Hits, pages, API calls, errors
- ◆ All **matching filters** (see filtering)
- ◆ **Custom counters** and top 20 of each **extracted parameter**

## 5.8 WHOIS

To know quickly who is behind a certain IP address, LogView Pro has a built-in WHOIS query. It is accessed via the **Display Menu** or by pressing **F9**.

```
% [whois.apnic.net]
% Whois data copyright terms    http://www.apnic.net/db/dbcopyright.html

% Information related to '223.204.0.0 - 223.205.255.255'

% Abuse contact for '223.204.0.0 - 223.205.255.255' is 'ipadmin@3bbmail.com'

inetnum:        223.204.0.0 - 223.205.255.255
netname:        TTBP-TH
descr:          Triple T Broadband Public Company Limited
country:        TH
admin-c:        TTBP1-AP
tech-c:         TTBP1-AP
abuse-c:        AT1597-AP
status:         ALLOCATED NON-PORTABLE
mnt-by:         MAINT-TTBP-TH
mnt-irt:        IRT-TTBP-TH
last-modified:  2021-12-22T04:23:17Z
source:        APNIC

irt:            IRT-TTBP-TH
address:        Jasmine International Public Company Limited,Pakkret Nonthaburi 11120
e-mail:         ipadmin@3bbmail.com
abuse-mailbox:  ipadmin@3bbmail.com
admin-c:        TTBP1-AP
tech-c:         TTBP1-AP
auth:           # Filtered
remarks:        ipadmin@3bbmail.com was validated on 2022-04-07
mnt-by:         MAINT-TTBP-TH
last-modified:  2022-09-02T09:36:30Z
source:        APNIC
```

WHOIS is an outdated standard and there is no standardized format. To make things worse, there are five registries worldwide, and each of them can use different representations for the results. That means, the built in WHOIS query works as follows:

- ◆ The WHOIS servers of each registry are queried one-by-one
- ◆ The first useful result (if the registry was responsible for the IP address) is returned

The WHOIS servers are queried in the order as defined in the **Global Settings**. It might make sense to change the order, depending on where the traffic is usually coming from.

## 5.9 RDAP

RDAP is the successor protocol to WHOIS and will replace it in the following years. It's a REST-like API, returning structured and standardized JSON data. The RDAP servers also should redirect queries to the responsible registries automatically. So only one server has to be configured.

The RDAP query is accessed via the **Display Menu**, or by pressing **F10**.

Address Range	223.204.0.0 - 223.205.255.255
Cidr Range(s)	223.204.0.0/15
Name	TTBP-TH
Type	ALLOCATED NON-PORTABLE
Country	TH
registration	22.12.2021 04:23:01
last changed	22.12.2021 04:23:17
TTBP1-AP	administrative technical
	Triple T Broadband PCL administrator Jasmine International Public Company Limited, Pakkret +6621008552 wanchai.ti@jasmine.com
IRT-TTBP-TH	abuse
	IRT-TTBP-TH Jasmine International Public Company Limited, Pakkret Nonthaburi ipadmin@3bbmail.com

In LogView Pro, the result is displayed in a text based and **simplified** form. RDAP queries are also used for other functions, for example sub-filter (“All in same IP range”).

The results

IP range or CIDR range can be copied and **used as filter** or used in the **Networks** list (see Global Settings)

## 5.10 Trace Route

Sometimes, the country information in the logfile, or provided by the country database is not correct. Even the **WHOIS** or **RDAP** entries could be **fake**.

To know where a certain IP address is located in reality, performing a trace-route can give a hint. Here an example: Cloudflare and the country database both locate the IP address in **Germany**, while WHOIS and RDAP both show **Ireland** as country. A trace-route shows a different picture:

```
Trace Route for 191.101.95.203 DE (Germany)

 1  17 ms  ZZ  192.168.1.1
 2  22 ms  ZZ  10.226.42.1
 3  14 ms  CH  unused.senselan.ch [83.222.128.45]
 4  *      Timed out
 5  20 ms  NL  ch-nax01a-rc1-ae-12-0.aorta.net [84.116.140.13]
 6  67 ms  NL  ch-zrh03a-rc2-ae-3-0.aorta.net [84.116.130.2]
 7  35 ms  NL  de-fra11b-rc1-ae-7-0.aorta.net [84.116.132.178]
 8  69 ms  NL  de-fra02a-ri1-ae-48-0.aorta.net [84.116.130.62]
 9  *      Timed out
10  *      Timed out
11  67 ms  ES  vlan3901.core2.fra4.de.m247.ro [37.120.220.91]
12  20 ms  RO  146.70.0.61
13  37 ms  GB  vlan2925.as03.fra4.de.m247.ro [83.97.21.17]
14  20 ms  DE  191.101.95.203
```

We can **ignore the title and the last line** (this is what the country DB shows). But the data packets seem to go travel in direction of **Romania**.

The built in trace-route tool is **much faster than the command line tool**, because the timeouts are shorter and the pings are executed in parallel.

Some Notes:

- ◆ The country information for the title and the single hops is taken from the **country database**, never from the log file.
- ◆ Because the query is still relatively slow compared to the other view modes, the view is automatically switched back to the session detail view when another session is clicked.

Even if this tool is very specialized and rarely used, it can be helpful in certain special cases.

## 5.11 IP Address List

This is a summary of all IP addresses of all the currently active (filtered) sessions. The list is accessed via the **Display Menu or Ctrl+F9**.

IP Address	Country	Sessions	Pages	API	Hits
113.53.122.149	TH	2	132	128	1427
1.46.203.45	TH	1	193	178	1345
49.49.201.121	TH	1	122	114	1335
124.120.228.171	TH	1	279	107	1264
182.232.81.217	TH	1	108	97	1160
118.172.244.136	TH	1	108	92	1144
202.151.5.93	TH	2	250	19	1135
1.4.192.41	TH	1	108	82	1022
49.230.198.115	TH	2	134	106	997
115.87.239.226	TH	3	105	81	977
1.47.235.233	TH	1	103	83	952
223.205.232.83	TH	2	194	31	945
58.8.173.254	TH	1	84	75	887
1.47.39.40	TH	2	124	84	798
61.19.219.152	TH	2	104	138	764
182.232.245.217	TH	3	104	171	720
182.232.69.115	TH	3	105	102	707
49.230.217.176	TH	1	168	21	697
161.246.37.196	TH	1	102	92	697

The results include the total summarized number of sessions, pages, API calls and hits for each IP Address.

The result can be sorted by clicking on the headers.

**Double clicking** a row creates a **sub-filter** instantly for that specific IP Address.

## 5.12 Useragent List

This is a summary of different user agents of all the currently active (filtered) sessions. The list is accessed via the **Display Menu or Ctrl+F10**.

Useragent	Sessions	Pages	API	Hits
Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/57.0.298...	161	712	795	7427
Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome...	116	402	399	4598
Mozilla/5.0 (iPhone; CPU iPhone OS 10_3_1 like Mac OS X) AppleWebKit/603.1.30 (KHTML...	89	339	348	3376
Mozilla/5.0 (iPhone; CPU iPhone OS 10_2_1 like Mac OS X) AppleWebKit/602.4.6 (KHTML, ...	56	327	281	3278
Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chr...	68	458	280	3259
Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chro...	80	288	277	3035
Mozilla/5.0 (Windows NT 5.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/49.0.262...	46	260	249	2510
Mozilla/5.0 (iPhone; CPU iPhone OS 10_3_1 like Mac OS X) AppleWebKit/602.1.50 (KHTML...	37	207	202	2305
Mozilla/5.0 (compatible; MJ12bot/v1.4.7; http://mj12bot.com/)	12	1374	0	1957
Mozilla/5.0 (Linux; Android 6.0; vivo 1601 Build/MRA58K) AppleWebKit/537.36 (KHTML, li...	18	259	166	1878
Mozilla/5.0 (Linux; Android 6.0.1; SM-J700F Build/MMB29K; wv) AppleWebKit/537.36 (KH...	28	212	164	1744
Mozilla/5.0 (compatible; AhrefsBot/5.2; http://ahrefs.com/robot/)	1508	1416	0	1607
Mozilla/5.0 (Linux; Android 6.0.1; SM-G610F Build/MMB29K) AppleWebKit/537.36 (KHTM...	11	122	118	1394
Mozilla/5.0 (Linux; Android 6.0.1; SM-A710F Build/MMB29K; wv) AppleWebKit/537.36 (KH...	8	138	93	1346
Mozilla/5.0 (Linux; Android 4.4.4; SM-T116NU Build/KTU84P) AppleWebKit/537.36 (KHTM...	1	193	178	1345
Mozilla/5.0 (Linux; U; Android 4.1.2; th-th; GT-I8552 Build/JZO54K) AppleWebKit/534.30 (K...	1	122	114	1335
Mozilla/5.0 (Windows NT 6.3; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome...	25	112	121	1300
Mozilla/5.0 (Linux; Android 5.1.1; F1f Build/LMY47V) AppleWebKit/537.36 (KHTML, like Ge...	10	127	101	1297
Mozilla/5.0 (Linux; Android 6.0; CAM-L21 Build/HUAWAICAM-L21) AppleWebKit/537.36 (...	12	123	137	1296

The results include the total summarized number of sessions, pages, API calls and hits for each different user agent.

The result can be sorted by clicking on the headers.

**Double clicking** in a row creates a **sub-filter** instantly for that specific user agent.

## 5.13 Country List

This is a summary of different countries of all the currently active (filtered) sessions. The list is accessed via the **Display Menu or Ctrl+F11**

ISO	Country	Sessions	Pages	API	Hits
TH	Thailand	3639	20115	17753	187898
US	United States	1257	985	241	3565
FR	France	1520	1670	1	2140
RU	Russian Federation	283	842	32	1442
DE	Germany	26	805	15	1334
SG	Singapore	9	55	45	605
LA	Lao People's Democratic Republic	24	54	37	528
IS	Iceland	4	20	16	289
KR	Korea, Republic of	7	13	7	207
AE	United Arab Emirates	2	15	14	177
IE	Ireland	88	5	1	106
JP	Japan	5	101	0	105
KZ	Kazakhstan	26	55	1	100
MM	Myanmar	4	6	3	88
CH	Switzerland	5	11	9	74

The results include the total summarized number of sessions, pages, API calls and hits for each different country.

The result can be sorted by clicking on the headers.

**Double clicking** in a row creates a **sub-filter** instantly for that specific country.



## 5.14 Referrer List

This is a summary of different referring domains of all the currently active (filtered) sessions. The list is accessed via the **Display Menu or Ctrl+F12**

Domain	Sessions	Pages	API	Hits
www.google.co.th	1149	8420	8326	84513
m.facebook.com	918	5390	3831	44792
www.kainetmotos.com	436	4703	4028	39716
(none)	4247	4663	455	13427
com.google.android.googlequicksearchbox	164	1042	1007	10672
www.facebook.com	28	191	208	2261
lm.facebook.com	8	103	89	879
www.google.com	10	91	83	662
www.google.la	2	52	41	486
r.search.yahoo.com	3	33	30	328
www.priceza.com	16	27	0	313
www.google.com.hk	4	22	19	245
www.google.co.nz	1	22	16	176
l.facebook.com	5	15	13	171
www.google.co.id	1	7	6	101

The results include the total summarized number of sessions, pages, API calls and hits for each referring domain.

The result can be sorted by clicking on the headers.

**Double clicking** in a row creates a **sub-filter** instantly for that specific referring domain.

## 5.15 Bot Scores List

If the log contains bot scores, a summary of the bot scores for all the currently active (filtered) sessions can be displayed. The list is accessed via the **Display Menu**

BotScore	IPs	Sessions	Pages	API	Hits	% Ses
0 - 9	324	343	1731	20879	28008	0.16
10 - 19	1076	1105	4870	78313	101648	0.52
20 - 29	2029	2106	8241	139432	180943	1
30 - 39	2907	3092	9909	193225	246992	1.46
40 - 49	3405	3679	11093	222150	284238	1.74
50 - 59	3419	3732	11088	232221	294587	1.76
60 - 69	4087	4454	13575	270541	343983	2.11
70 - 79	5150	5686	17021	351066	446069	2.69
80 - 89	9506	10972	29732	690810	863768	5.19
90 - 99	82504	176414	450472	10806630	13445299	83.38

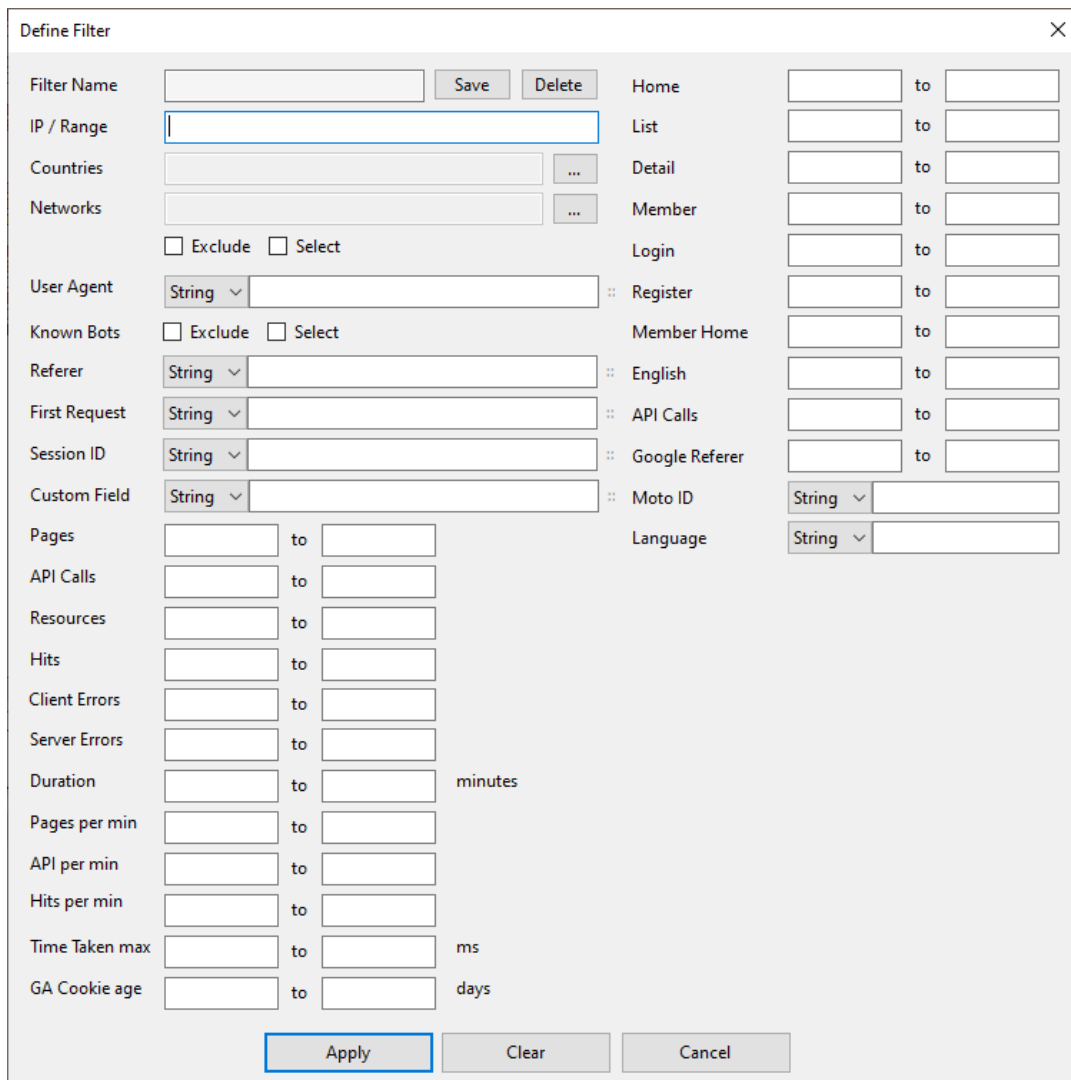
The result can be sorted by clicking on the headers.

**Double clicking** in a row creates a **sub-filter** instantly for that specific bot score range.

## 6 Filtering

### 6.1 Defining a Filter

As a day of web traffic can contain hundreds of thousands of sessions, it is necessary to filter it, so one can have a look at just a part of the traffic. The filter form is accessed via the **Filter menu** or by pressing **Ctrl+F**:



The filter form has two main parts:

- ◆ Left column: **Standard filter** criteria, all available by default
- ◆ Other columns to the right: Filter for **custom counters** and **parameters**

The form is dynamic, depending on how many counters and parameters are defined in the configuration.

These are the standard filter criteria:

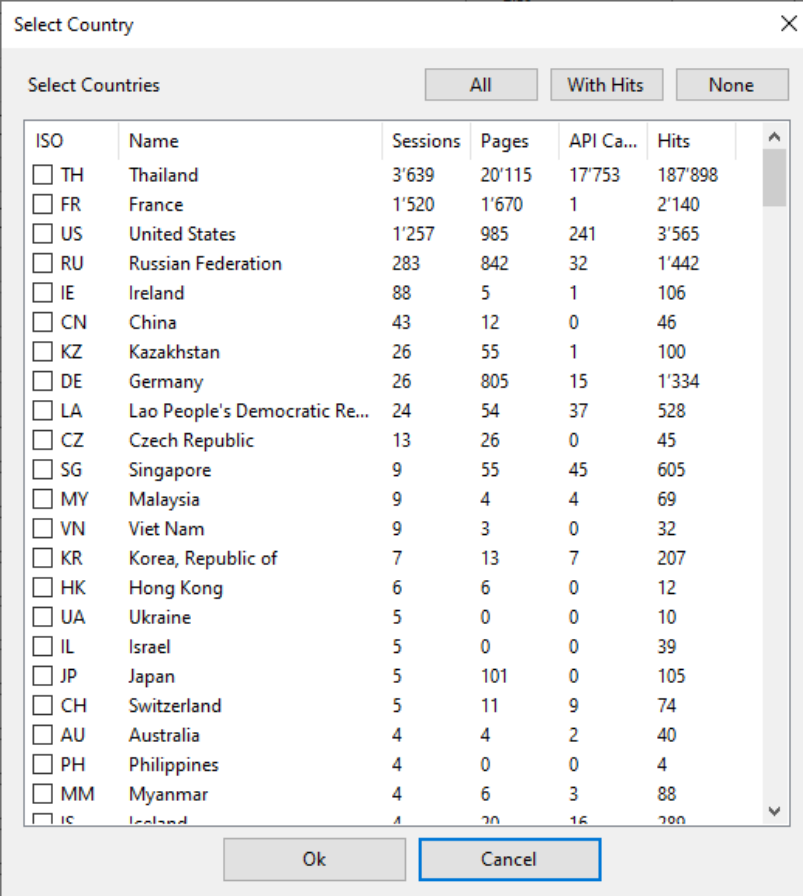
Filter Name	Filter name, if it is saved as a <b>filter preset</b>
IP / Range	<p>Single <b>IP address or entire range</b>. The following notations are possible:</p> <p>IPv4 address: 192.168.0.1</p> <p>IPv4 range: 192.168.0.1 - 192.168.255.255</p> <p>IPv4 CIDR: 178.197.128.0/17</p> <p>IPv6 address: 2a03:d000:4220:2528:e0a2:cfea:a7a2:812</p> <p>IPv6 range: 2001:db0:: - 2001:0dbf:ffff:ffff:ffff:ffff:ffff:ffff</p> <p>IPv6 CIDR: 2a03:d000:4200::/40</p> <p>To select all IPv6 addresses, just use the notation <b>::/0</b></p>
Countries	List of ISO-2-letter codes for countries that should be filtered for. These are selected in the <b>country selector form</b> (see below)
Networks Exclude / Select	<b>Networks</b> can be defined in the <b>Global Settings</b> . After selecting them in the <b>network selector form</b> (see below), they can be included or excluded. If no check box is checked, nothing is filtered.
User Agent	Search expression for a <b>specific useragent</b> . Example: Mozilla/5.0 (Linux; Android *)
Known Bots Exclude / Select	<p>The known (good) bots are defined in the <b>Global Settings</b>. These can be included or excluded.</p> <p><b>Note:</b> There is always the risk that a bot uses the user agent of a legitimate search engine as camouflage. But practice has shown that this is relatively rare, but it's always worth keeping an eye on that.</p>
Referer	<b>Referer header</b> that is detected for the session. Usually it is taken from the <b>first page request</b> within the session. Referer-headers from API-calls and resource-requests are ignored.
First Request	The <b>very first</b> request within the session.
Session ID	The <b>session ID</b> or session cookie. In case of <b>Google Analytics</b> cookies, the first two digits are replaced by an x. Example: GAx.x.1372917134.1641682372
Custom Field	Any content of any <b>Custom Field</b> (see Configuration). Note that it's not possible to search for empty (NULL) values.
Pages	Number of <b>page requests</b> within the session (not including redirects)
API Calls	Number of <b>API calls</b> within the session
Resources	Number of requests that are <b>neither pages or API calls</b>
Hits	Number of <b>total requests</b> (also called hits)
Client Errors	Number of client errors (HTTP status 4xx)
Server Errors	Number or server errors (HTTP status 5xx)
Duration (minutes)	Duration of the entire session

Pages per min	Speed of the page requests, measured in pages per minute
API per min	Speed of the API calls, measured in API calls per minute
Hits per min	Speed of overall requests (hits), measured also per minute
Time taken max	Longest execution time of a request (time taken) of all the requests within a session
GA Cookie age	If Google Analytics cookies are used for session identification, the age of the cookie is part of the value and can be easily decoded. So it's possible to search for the <b>age of the cookie</b> .

The buttons have the following functions:

Apply	Apply the filter to the loaded traffic.
Clear	<b>Clear</b> the current filter.
Cancel	<b>Cancel</b> editing of the filter. Nothing is changed or applied.
Save	<p><b>Save the filter as a preset</b> (see chapter about presets). To save a filter, it needs a name.</p> <p>If the name is changed, a new filter is saved without changing the old one. This way it is possible to create new filters based on another one without changing the original.</p> <p>To change the name of a filter, it has to be saved under the new name, and the original one has to be deleted.</p>
Delete	<b>Delete the filter</b> with the name currently set. But the filter can stay active in the memory

The **country selector form** allows easy selection of countries to filter for:



ISO	Name	Sessions	Pages	API Ca...	Hits
<input type="checkbox"/>	TH Thailand	3'639	20'115	17'753	187'898
<input type="checkbox"/>	FR France	1'520	1'670	1	2'140
<input type="checkbox"/>	US United States	1'257	985	241	3'565
<input type="checkbox"/>	RU Russian Federation	283	842	32	1'442
<input type="checkbox"/>	IE Ireland	88	5	1	106
<input type="checkbox"/>	CN China	43	12	0	46
<input type="checkbox"/>	KZ Kazakhstan	26	55	1	100
<input type="checkbox"/>	DE Germany	26	805	15	1'334
<input type="checkbox"/>	LA Lao People's Democratic Re...	24	54	37	528
<input type="checkbox"/>	CZ Czech Republic	13	26	0	45
<input type="checkbox"/>	SG Singapore	9	55	45	605
<input type="checkbox"/>	MY Malaysia	9	4	4	69
<input type="checkbox"/>	VN Viet Nam	9	3	0	32
<input type="checkbox"/>	KR Korea, Republic of	7	13	7	207
<input type="checkbox"/>	HK Hong Kong	6	6	0	12
<input type="checkbox"/>	UA Ukraine	5	0	0	10
<input type="checkbox"/>	IL Israel	5	0	0	39
<input type="checkbox"/>	JP Japan	5	101	0	105
<input type="checkbox"/>	CH Switzerland	5	11	9	74
<input type="checkbox"/>	AU Australia	4	4	2	40
<input type="checkbox"/>	PH Philippines	4	0	0	4
<input type="checkbox"/>	MM Myanmar	4	6	3	88
<input type="checkbox"/>	IS Iceland	4	20	16	200

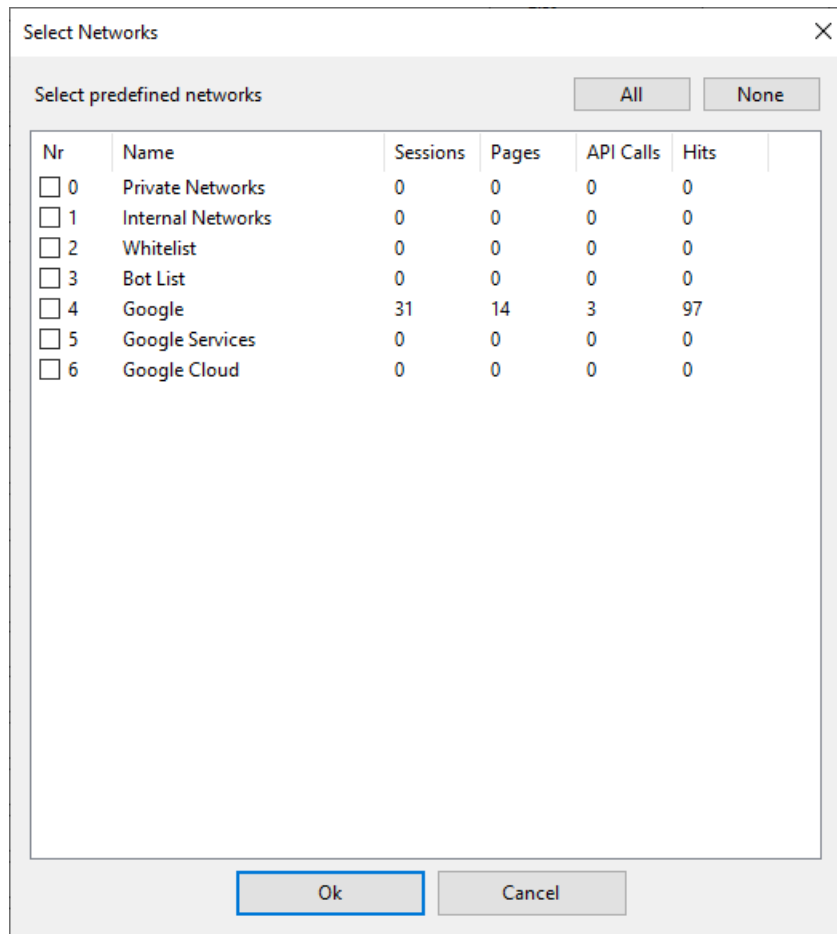
The number for sessions, pages, API calls and hits represent the numbers for the entire traffic, not just the currently filtered sessions.

The columns can be sorted.

- ◆ **All** selects all the countries
- ◆ **With Hits** only checks the countries that in fact have traffic
- ◆ **None** deselects all countries

To **remove** the country filter completely, so no filtering for countries is done at all, just select **None** (this is better for performance than selecting all the countries).

In the **network selector form** the networks defined in the **Global Settings** can be selected for filtering:



Nr	Name	Sessions	Pages	API Calls	Hits
<input type="checkbox"/> 0	Private Networks	0	0	0	0
<input type="checkbox"/> 1	Internal Networks	0	0	0	0
<input type="checkbox"/> 2	Whitelist	0	0	0	0
<input type="checkbox"/> 3	Bot List	0	0	0	0
<input type="checkbox"/> 4	Google	31	14	3	97
<input type="checkbox"/> 5	Google Services	0	0	0	0
<input type="checkbox"/> 6	Google Cloud	0	0	0	0

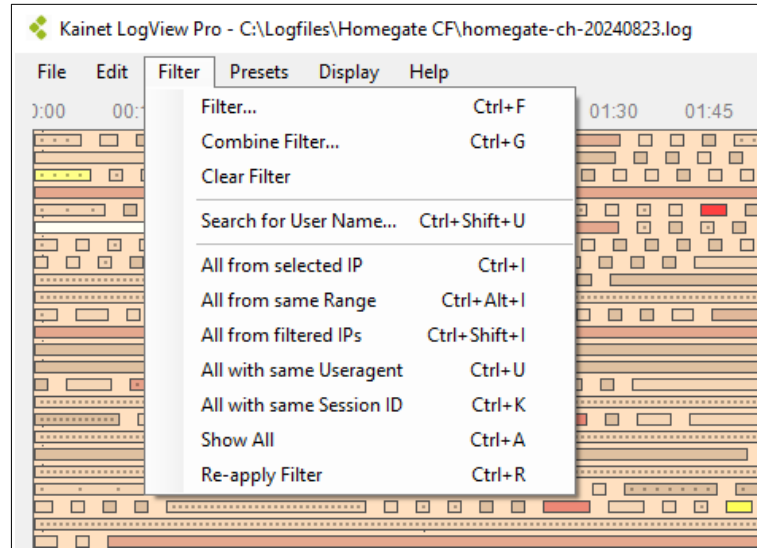
The number for sessions, pages, API calls and hits represent the numbers for the entire traffic, not just the currently filtered sessions.

Please note that for larger logfiles, the calculation of the counters would take too long, so after a few seconds, the counting is cancelled and only an incomplete count is displayed.

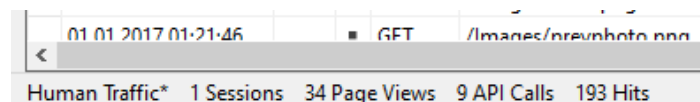
To **remove** the network filter completely, so no filtering for networks is done at all, just select **None**.

## 6.2 Sub-Filter

So-called sub-filters can be applied for a **quick lookup** for certain traffic similar to the currently active or selected sessions. The original filter remains in the background and can be re-applied. The sub-filters are activated in the **Filter Menu**:



All from selected IP	A session has to be selected. Then a temporary filter containing <b>only sessions with the same IP address</b> is applied.
All from same Range	A RDAP lookup is executed and <b>all the sessions in the same IP range</b> as the currently selected one are displayed.
All from filtered IPs	The current selection is <b>expanded to all the sessions that share IP addresses</b> with the currently filtered (active) ones.
All with same Useragent	All sessions with the <b>same user agent</b> (session Cookie) as the currently selected.
All with same Session ID	All sessions with the <b>same session ID</b> as the currently selected. Usually it's the same user.
Show All	Temporarily <b>show the entire traffic</b> while keeping the current filter in the background
Re-apply Filter	Re-apply and go back to the original filter



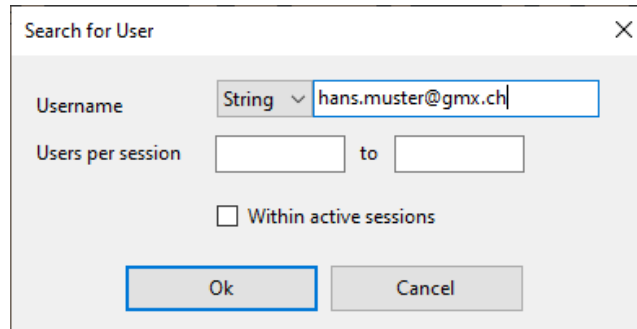
A star \* indicates that a sub-filter is active and the original filter can be restored.



## 6.3 Search for user names

If the current authenticated user is logged, it is possible to search for it. This is a special sub-filter and not part of a normal filter.

The following window is displayed:



If a authenticated session was selected, the current user name is suggested by default. These are the options:

Username	Search Expression to search for a a user name
Users per session	If a user has logged in into several user accounts during a single session, he can be identified with this search. This can be useful to find scammers who acquired user credentials with a phishing attack
Within active sessions	The user search can be performed either on the entire loaded traffic or just the currently filtered users

With re-apply filter, the previously active filter can be called back.

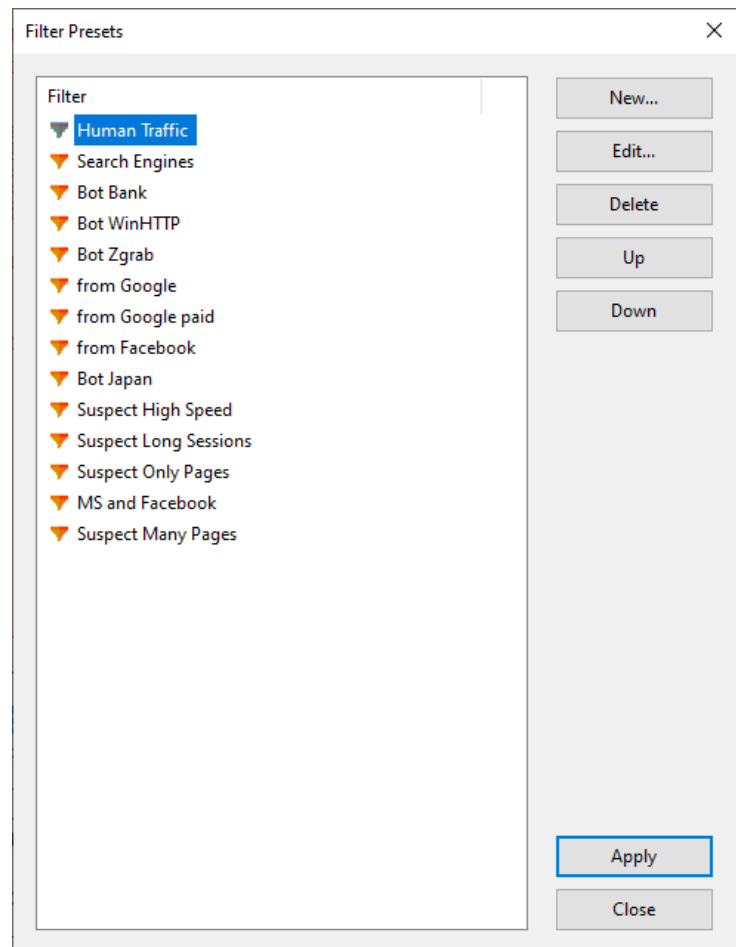
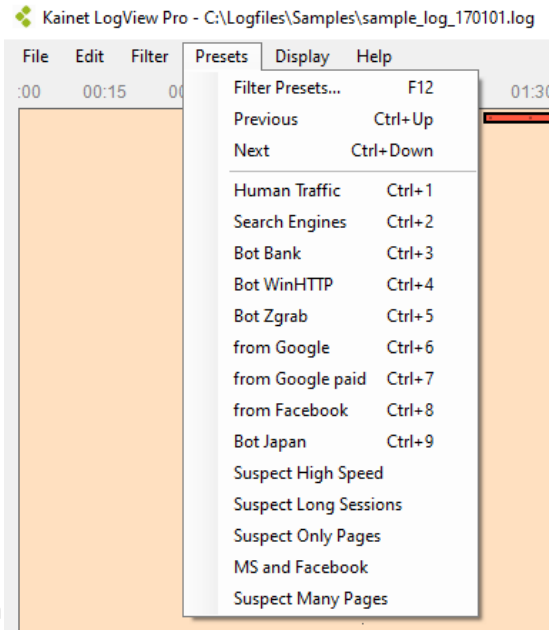
## 6.4 Working with Filter Presets

When a filter is saved, it will be available as preset. The presets can be accessed in two ways:

- ◆ The first 30 presets are directly added to the **Presets Menu**
- ◆ The preset manager can be accessed via the **Presets Menu** or by pressing **F12**

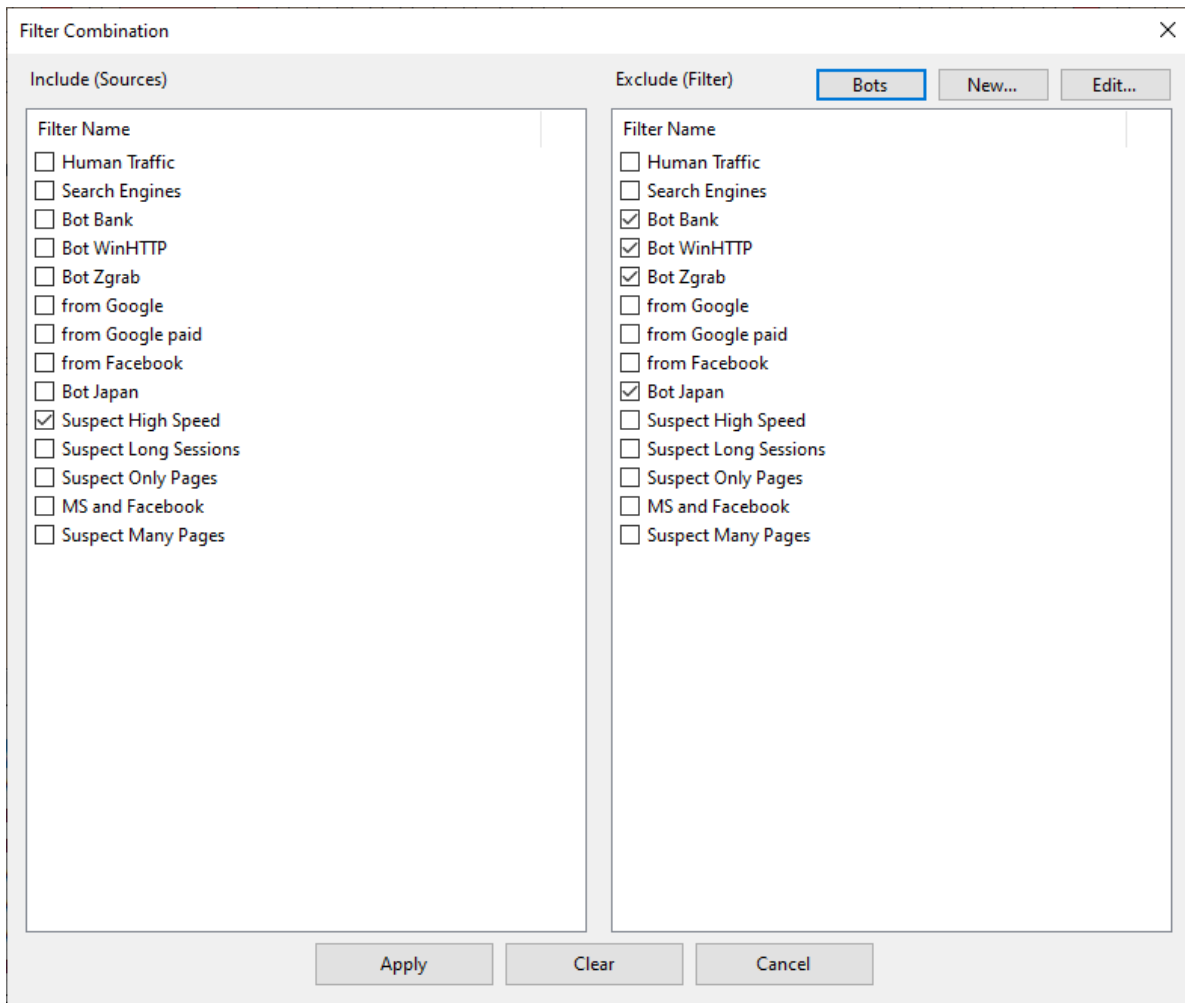
In the preset manager, filters can be **added**, **deleted** or **edited**. The order of the presets can also be altered, by using **Up** or **Down**.

To quickly switch between filter presets and apply them without calling the menu or the presets manager, the shortcuts **Ctrl-Down** and **Ctrl-Up** can be used in the main window.



## 6.5 Filter Combinations

Filter combinations are useful if there are many different bots currently active on a website. Do differentiate them, and to find new ones which are not identified yet, there is the possibility to combine filter presets:



Filters on the **left side** will **include** all the sessions that match any one of the selected filters.

Filters on the **right side** will **exclude** all the sessions that match any of the selected filters, among the included ones from left side.

So, a session will be displayed if it matches any selected filter on the left, but no selected filter on the right.

The button **Bots..** automatically selects all the filters that contain “bot” in their name. This is helpful if there is already a long list of filters for specific bots that should be excluded.

## 7 Copy commands

The standard copy command **Ctrl-C** works a bit differently depending on which view is currently active:

Traffic Statistics	A bitmap containing the graph is copied
Summary Session Summary WHOIS RDAP	The selected text, or (if nothing is selected) the entire summary is copied to the clipboard as text
Session Details Session Pages IP Address List User Agents List Country List Referrer List	The selected cells are copied to the clipboard as text

There are special copy commands, all to be accessed via the **Edit Menu**. For those, a session has to be selected:

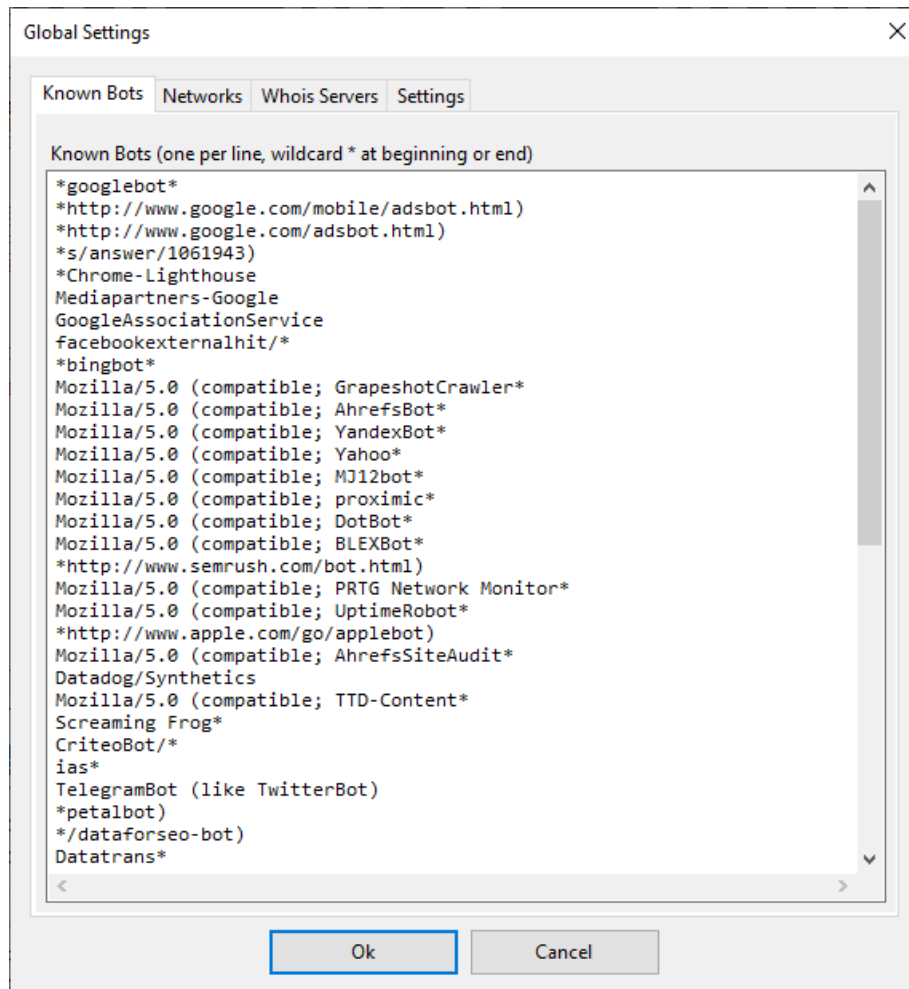
Copy IP Address	The IP address of the currently selected session is copied
Copy IP Range	A RDAP query is executed and the ip address range that is found for the currently selected session is copied
Copy User Agent	The user agent of the currently selected session is copied
Copy Session ID	The session ID of the currently selected session is copied

Main purpose of these shortcuts is to make it easier and quicker to copy the information needed to **create a new filter**.

## 8 Settings

### 8.1 Known Bots

This list contains the patterns matching the user agents of known (good) bots. These can be **included or excluded in the filters**.



Usage:

- ◆ One line per user agent
- ◆ The patterns can contain \* as wildcard at the beginning or end, but not in the middle

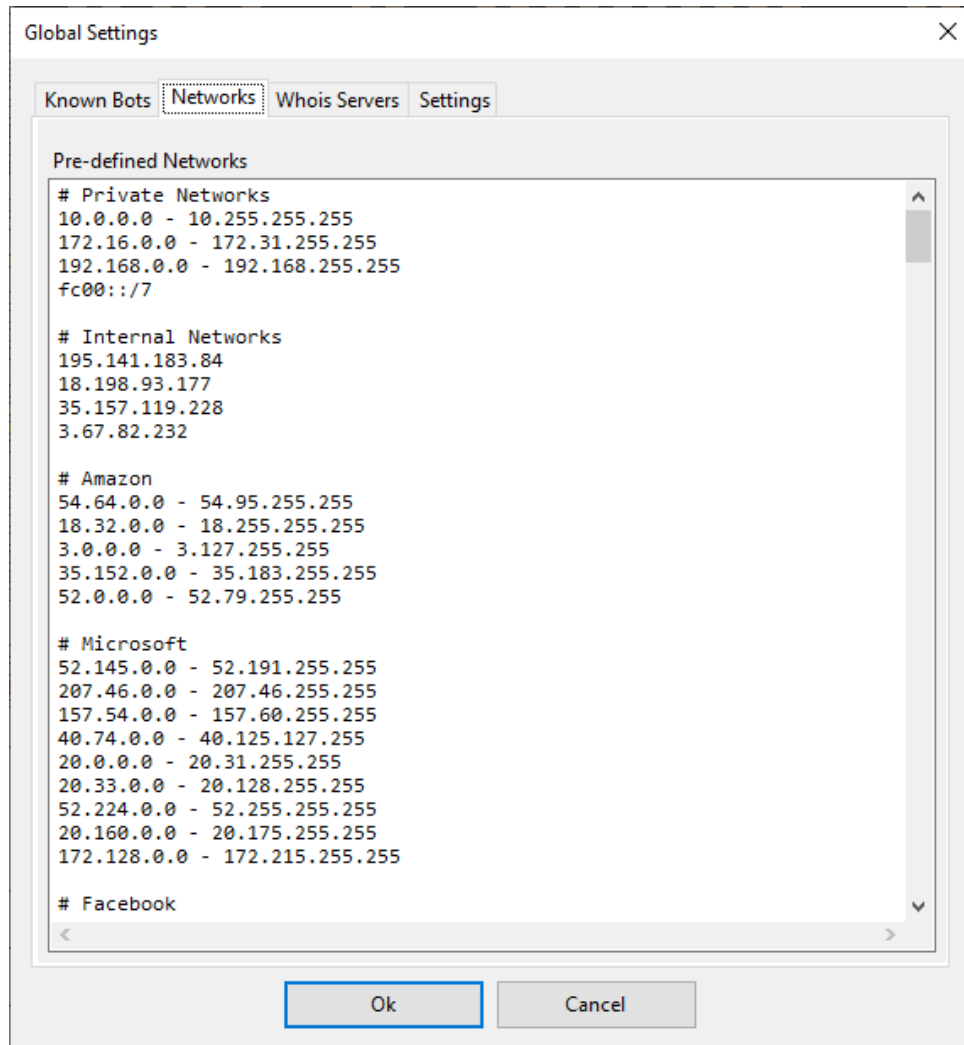
Notes:

- ◆ The performance is better if the wildcard \* is only used at the beginning or the end, but not both.
- ◆ Bad bots can fake the user agent and use one of a tolerated good bot (for example googlebot). While this is rare, is always have to be remembered.

## 8.2 Networks

Kainet LogView Pro does not include a network database, as this would cause additional cost, including a monthly subscription. A big database would also affect the performance negatively.

But it is possible to define an own list of networks of interest, which can be used during filtering:



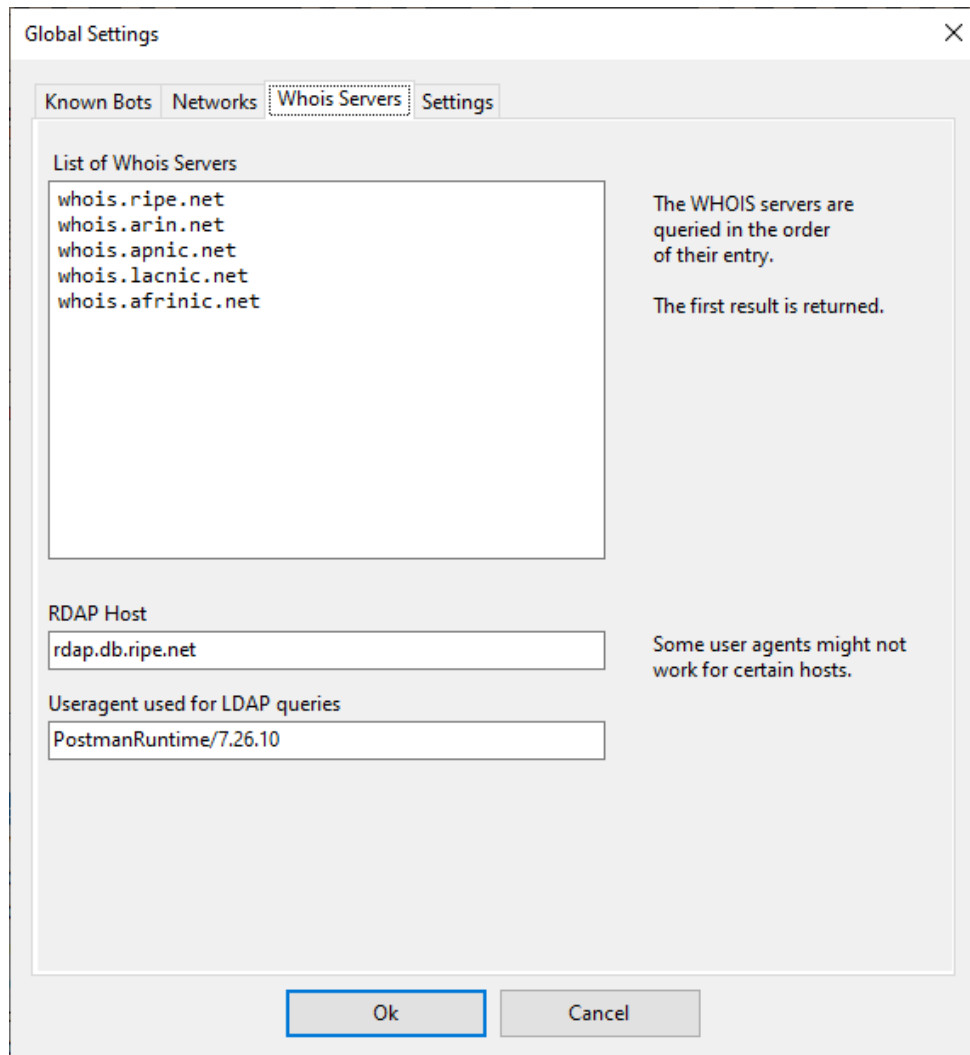
The syntax is as follows:

- ◆ **Hashtag + Network name** starts a network section
- ◆ Followed by a **list of single IP addresses** or **network ranges**. The **CIDR** notation can be used, but in case of IPv4 it will be converted to a “normal” range.

**Important:** Do not change the name of a pre-defined network once it is used by a filter preset.

## 8.3 WHOIS / RDAP settings

The servers used for WHOIS and RDAP settings are defined here:



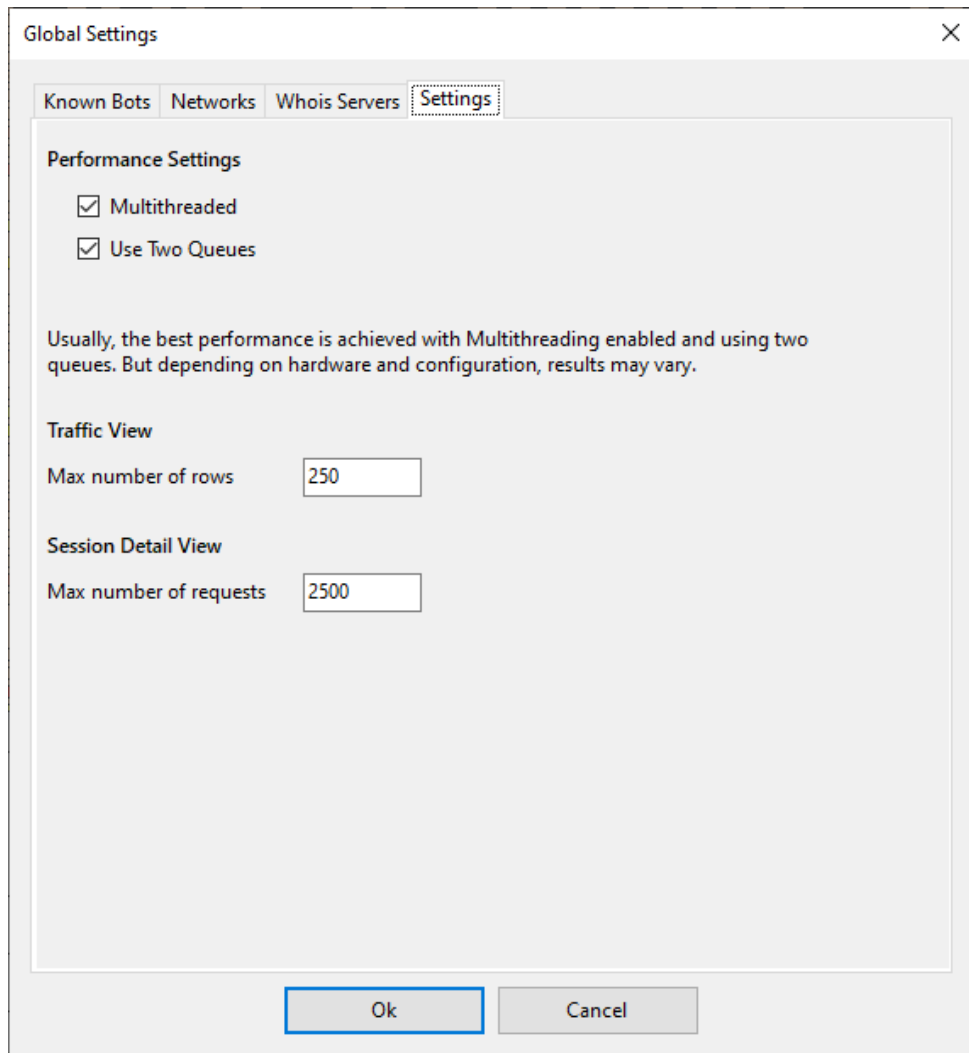
Some notes:

- ◆ The optimal order of the WHOIS servers depends on where most of the traffic is coming from, as every server has to be queried until a useful result is returned
- ◆ For RDAP only one host must be configured, as the protocol requires requests to be redirected to the responsible registry
- ◆ Some RDAP hosts don't accept any user agent (for example rdap.org). That's why the default user agent is Postman, even as we don't use Postman for queries at all.

We are using caching for RDAP queries, so we avoid multiple queries for the same IP address. Nevertheless we ask our users to use these features responsibly in order to avoid being blocked.

## 8.4 Other Settings

All other application settings are found in this tab:



Currently only the multithreading options are found here:

<p>Multithreaded (recommended: ON)</p>	<p>Multithreading is enabled. The benefit of using multithreading depends on the system (number of cores, memory throughput etc.) and on the configuration for a website.</p>
<p>Use Two Queues (recommended: OFF in most cases)</p>	<p>Loading of logfiles consists of three steps. Enabling this option lets the last step run in one or more separate threads, depending on the number of cores.  For typical systems and configurations there is no benefit in enabling this, rather the opposite.  But if a configuration contains a lot of custom counters and parameters, and if the computer has good memory performance, a certain improvement is possible.</p>



---

Traffic View / Max number of rows	With large logfiles, applying a filter or subfilter with a lot of sessions can take a long time. So the maximum number of rows in the main traffic display can be limited. In other words, this limits the number of simultaneous sessions displayed.
Session Detail View / Max number of requests	Limits the number of requests that are loaded into the session detail.

---

## 9 Strategies

There are many strategies to identify malicious traffic. For the moment we just list the most important steps recommended. This chapter will be expanded in the future:

- ◆ Create a configuration that contains **counters** which cover most of the pages and API calls typically used by normal users, but also by bots
- ◆ Especially counters for requests that return valuable information (such as object detail views) are important.
- ◆ On the other hand, counters for requests that are typically NOT requested by bots, can be helpful when differentiating human traffic from bots
- ◆ Create a **filter** for traffic **excluding internal traffic** (caused by own systems, such as monitoring) and also **excluding legitimate search engines**. This filter will be the basis for more specific filters
- ◆ A filter that includes mostly human traffic only can be helpful
- ◆ Learn how **human traffic** and sessions by human users **usually look**
- ◆ Create special filters for different kind of **suspicious** traffic: Very long sessions, sessions with very high numbers of certain requests, many pages/ api calls per minute etc.
- ◆ Use the **Traffic Statistics** view and look out for **suspicious peaks** that don't follow the normal daily traffic distribution.
- ◆ Try to **isolate the traffic** that causes these peaks by modifying the filters.
- ◆ Look at **individual sessions** that are part of this suspicious traffic using the **Session Details** and the **Session Pages** view. This will confirm if it's really a bot or not.
- ◆ The **Session Summary** will be helpful when defining specific rules for individual bots
- ◆ Once bots are clearly identified, define filters that **isolate** these **single bots** as accurately as possible. In many cases it will not be possible to be one hundred percent accurate.
- ◆ Use the filter **combination form**: On the left side include one or more filters for suspicious traffic, on the right side exclude bots that are already identified. This way yet undetected bots can be found
- ◆ Repeat these steps until after exclusion of detected bots, **no more suspicious peaks** are visible
- ◆

LogView Pro offers a lot of functions, so depending on the practical case, many different strategies can be appropriate.

## 10 Useful information

### 10.1 Logging basics

#### 10.1.1 W3C Logfile fields / columns

<i>W3C</i>	<i>Name</i>	<i>Description</i>	
<b>date</b>	Date	The date on which the activity occurred.	**
<b>time</b>	Time	The time the activity occurred.	***
<b>c-ip</b>	Client IP Address	The IP address of the client that accessed your server.	***
<b>cs-username</b>	Username	The name of the user who accessed your server, if he was not connected anonymously	**
<b>s-sitename</b>	Service Name	The Internet service that was running on the client computer.	
<b>s-computername</b>	Server Name	The name of the server on which the log entry was generated.	
<b>s-ip</b>	Server IP	The IP address of the server on which the log entry was generated.	
<b>s-port</b>	Server Port	The port number the client is connected to. The standard port for HTTP is port 80.	
<b>cs-method</b>	Method	The action the client was trying to perform, usually GET or POST (see below)	**
<b>cs-uri-stem</b>	URI Stem	The resource accessed: for example, an HTML page, a CGI program, a script, an image, a media file.	***
<b>cs-uri-query</b>	URI Query	The query string, if any. The query string is separated from the Stem part of the URL by a question mark. The query string usually consists of one or more name-value pairs (see below)	***
<b>sc-status</b>	HTTP Status	The status of the action, in HTTP terms (see below)	**
<b>sc-win32-status</b>	Win32 Status	The status of the action, in terms used by Windows.	
<b>sc-bytes</b>	Bytes Sent	The number of bytes sent by the server.	*
<b>cs-bytes</b>	Bytes Received	The number of bytes received by the server.	
<b>time-taken</b>	Time Taken	The length of time the action took.	*
<b>cs-version</b>	Protocol Version	The protocol (HTTP, FTP) version used by the	

		client. For HTTP this will be either HTTP 1.0 or HTTP 1.1.	
<b>cs-host</b>	Host	The hostname requested in the url. As there is possible to run several web sites on the same IP address, the host name is used to make the differentiation	*
<b>cs(User-Agent)</b>	User Agent	The browser used on the client. It contains information about the browser manufacturer, the browser version and the operating system. Very often the user agent is manipulated or fake.	***
<b>cs(Cookie)</b>	Cookie	The cookies sent from the browser to the server, if any. The cookie which identifies a user session is found here.	*
<b>cs(Referer)</b>	Referer	The full url that directed the user to the current site.	**

\*\*\* Required by LogView Pro

\*\* Useful / Recommended

\* Supported und used by certain features

#### Non-standard fields supported by LogViewPro

<b>X-Forwarded-For</b>	Original IP address	If the servers reside behind a reverse proxy, this field contains a list of IP addresses that indicate the path the request took
<b>X-BotScore</b>	Bot Score	Bot score provided by a WAF or CDN
<b>X-Cookie</b>	Session ID	Session ID as separate field
<b>X-Country</b>	Country	ISO code of country of client ip address

The names of these fields can vary and **can be configured** in LogView Pro.

## 10.1.2 Methods

There are different HTTP request methods:

<b>GET</b>	By far the most common method used to request for a specified URL. When using GET, all the values sent to the server are part of the query string.
<b>HEAD</b>	Identical to GET, except that the page content is not returned; just the headers are. Useful for retrieving meta-information.
<b>POST</b>	Similar to GET, except that a message body, typically containing key-value pairs from an HTML form submission, is included in the request data, not as part of the URL. While all the get parameters can be found in the logfile, the post data is not logged. Typically POST is used for complex forms.
<b>TRACE</b>	Echoes back the received request, so that a client can see what intermediate servers are adding or changing in the request.
<b>OPTIONS</b>	Returns the HTTP methods that the server supports. This can be used to check the functionality of a web server.
<b>CONNECT</b>	For use with a proxy that can change to being an SSL tunnel
<b>PUT</b>	Used for uploading files to a specified URI on a web-server.
<b>DELETE</b>	Rarely implemented, deletes a resource (i.e. a file).

### 10.1.3 Query Strings

When data is sent to the server as part of the URL, then this is done using query strings. These can contain these special characters:

- ? Separates the resource from the query string
- & Delimiter between name-value pairs
- = Delimiter between name and value
- + Replacement for space character in query string (for details look for URL Encoding)
- %nn Encoded character

Example query string:

```
make=RENAULT&modellike=M%C3%A9gane&submit=Search+Car
```

Post data is encoded the same way, but not logged. Submitting a GET-Form also generates a url with a query string

## 10.1.4 Status Codes

The status code is returned by the server as result of each request:

### *Code Definition*

#### **1xx Informational**

- 100 CONTINUE - the client should continue with request.
- 101 SWITCHING PROTOCOLS - the server will switch protocols as necessary.

#### **2xx Successful**

- 200 OK - the request was fulfilled.
- 201 CREATED - following a *POST* command.
- 202 ACCEPTED - accepted for processing, but processing is not completed.
- 203 NON-AUTHORITATIVE INFORMATION - the returned metainformation is not the definitive set from the original server.
- 204 NO CONTENT - request received but no information exists to send back.
- 205 RESET CONTENT - the server has fulfilled the request and the user agent should reset the document view.
- 206 PARTIAL CONTENT - the server has fulfilled the partial GET request.

#### **3xx Redirection**

- 300 MULTIPLE CHOICES - the requested resource has many representations.
- 301 MOVED PERMANENTLY - the data requested has a new location and the change is permanent.
- 302 FOUND - the data requested has a different URL temporarily.
- 303 SEE OTHER - a suggestion for the client to try another location.
- 304 NOT MODIFIED - the document has not been modified as expected.
- 305 USE PROXY - The requested resource must be accessed through the specified proxy.
- 306 UNUSED
- 307 TEMPORARY REDIRECT - the requested data resides temporarily at a new location.

#### **4xx Client Errors**

- 400 BAD REQUEST - syntax problem in the request or it could not be satisfied.
- 401 UNAUTHORIZED - the client is not authorized to access data.
- 402 PAYMENT REQUIRED - indicates a charging scheme is in effect.
- 403 FORBIDDEN - access not required even with authorization.
- 404 NOT FOUND - server could not find the given resource.
- 405 METHOD NOT ALLOWED
- 406 NOT ACCEPTABLE
- 407 PROXY AUTHENTICATION REQUIRED - the client must first authenticate with the proxy for access.
- 408 REQUEST TIMEOUT - the client did not produce a request within the time the server was

prepared to wait.

409 CONFLICT - the request could not be completed due to a conflict with the current state of the resource.

410 GONE - the requested resource is no longer available.

411 LENGTH REQUIRED - the server refused to accept the request without a defined *Content Length*.

412 PRECONDITION FAILED

413 REQUESTED ENTITY TOO LARGE - the server is refusing to process a request because it is larger than the server is willing or able to process.

414 REQUEST-URI TOO LONG - the server is refusing to process a request because the *URI* is longer than the server is willing or able to process.

415 UNSUPPORTED MEDIA TYPE - requested resource format is not supported.

416 REQUESTED RANGE NOT SATISFIABLE

417 EXPECTATION FAILED

#### **5xx Server Errors**

500 INTERNAL ERROR - the server could not fulfill the request because of an unexpected condition.

501 NOT IMPLEMENTED - the sever does not support the facility requested.

502 BAD GATEWAY - received an invalid response from an upstream sever.

503 SERVICE UNAVAILABLE - the server is currently unable to handle a request.

504 GATEWAY TIMEOUT - The server, acting as a gateway/proxy, did not receive a timely response from an upstream server.

505 HTTP VERSION NOT SUPPORTED

### 10.1.5 Cookies

A HTTP magic cookie (usually called simply a cookie) is a packet of information sent by a server to a World Wide Web browser and then sent back by the browser each time it accesses that server (but not to other servers).

Cookies are often used to identify user sessions, because it is common that many computers share the same IP address towards the internet and there is no other way to differentiate them.



### 10.1.6 Sessions (Visits)

Every request is logged separately in the logfiles in chronological order, just as they get in. As a result, the requests of different users that are accessing a web server simultaneously, are completely mixed up.

LogViewPro tries to get these requests together again build sessions (or visits, as you like) by identifying the originator by some characteristics:

- IP Address
- User Agent
- Session Cookie (if available)

As the session cookie is not always available, and as there are very homogenous networks where many identical computers share the same IP Address, it is possible, that more than one session are counted and displayed as a single one. This is a problem that all the logfile analysis tools have to fight with.

### 10.1.7 DNS

The Domain Name System is a system that stores information about hostnames and domain names in a type of distributed database on the Internet. Of the many types of information that can be stored, most importantly it provides a physical location (IP address) for each domain name.

### 10.1.8 WHOIS

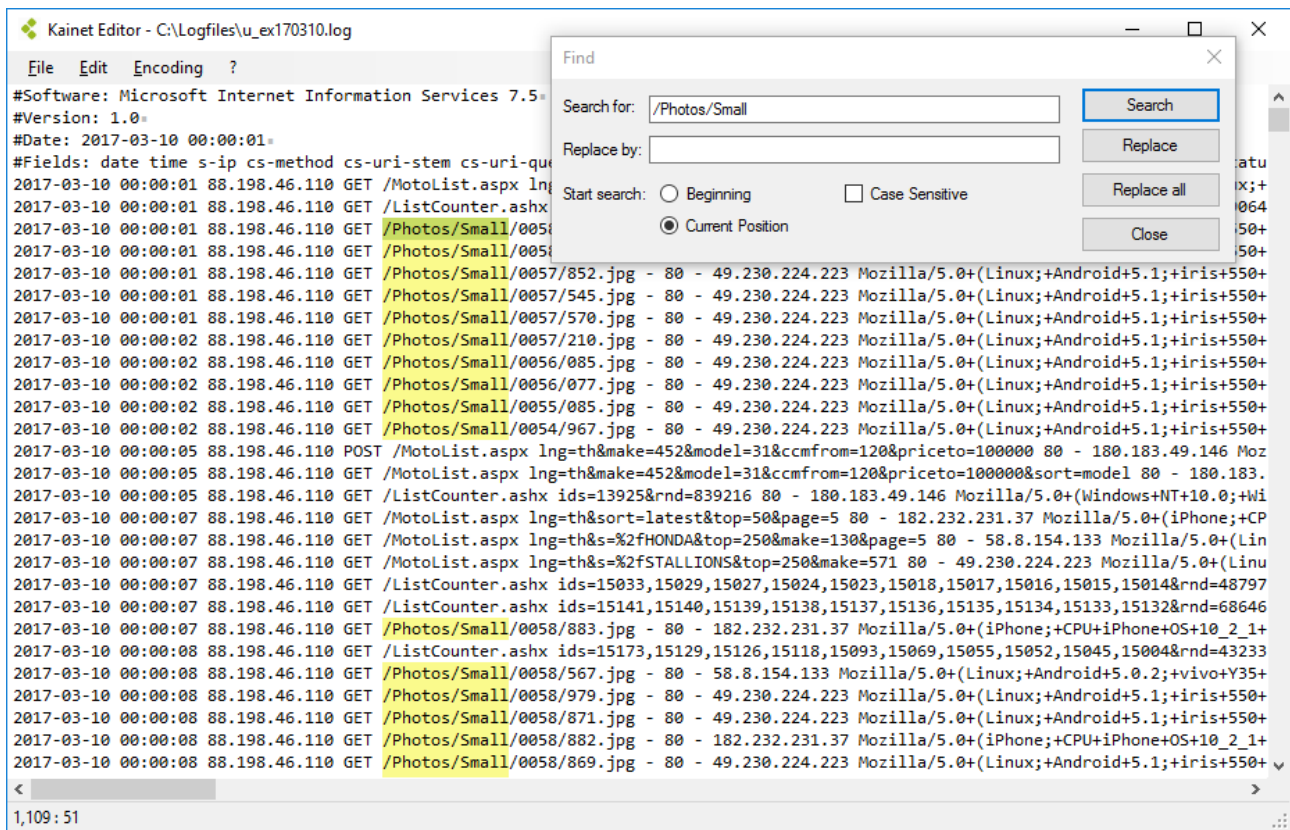
WHOIS is a TCP-based protocol which is widely used for querying a database in order to determine the owner of a domain name, an IP address, or an autonomous system number, on the Internet.

# 11 Other related software

## 11.1 Kainet Editor

If you want to have a look at your log files, for example to see what fields are logged or what format it is, you need an editor or viewer which is able to open huge files without loading them completely into memory.

Kainet editor not only lets you open huge files, you can also do changes or replacements in files of any size:



The screenshot shows the Kainet Editor application window titled "Kainet Editor - C:\Logfiles\u\_ex170310.log". The main text area displays a log file with the following content:

```

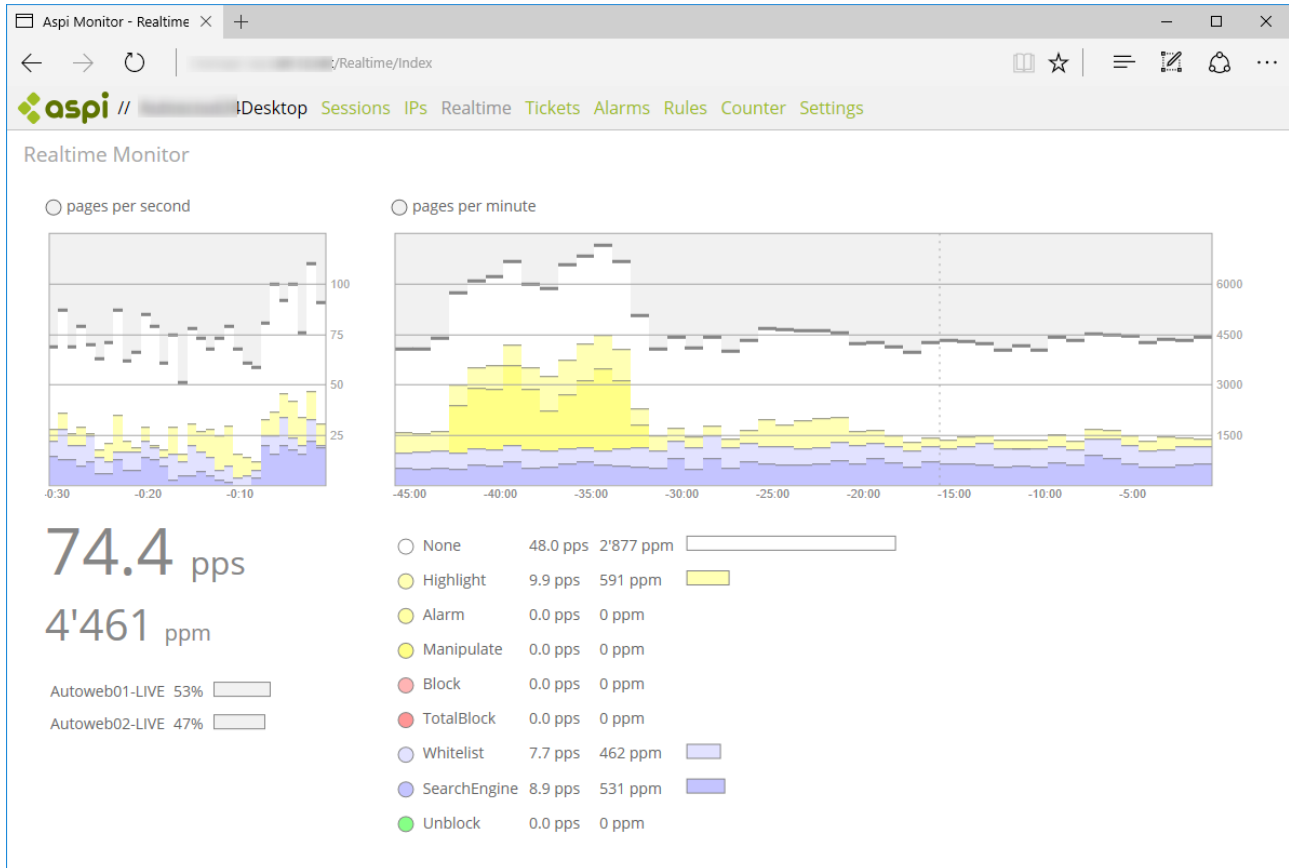
#Software: Microsoft Internet Information Services 7.5
#Version: 1.0
#Date: 2017-03-10 00:00:01
#Fields: date time s-ip cs-method cs-uri-stem cs-uri-quer
2017-03-10 00:00:01 88.198.46.110 GET /MotoList.aspx lng=th&make=452&model=31&ccmfrom=120&priceto=100000&sort=model 80 - 180.183.49.146 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:01 88.198.46.110 GET /ListCounter.ashx ids=13925&rnd=839216 80 - 180.183.49.146 Mozilla/5.0+(Windows+NT+10.0;+Wi
2017-03-10 00:00:01 88.198.46.110 GET /Photos/Small/0055/852.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:01 88.198.46.110 GET /Photos/Small/0057/545.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:01 88.198.46.110 GET /Photos/Small/0057/570.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:02 88.198.46.110 GET /Photos/Small/0057/210.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:02 88.198.46.110 GET /Photos/Small/0056/085.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:02 88.198.46.110 GET /Photos/Small/0056/077.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:02 88.198.46.110 GET /Photos/Small/0055/085.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:02 88.198.46.110 GET /Photos/Small/0054/967.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:05 88.198.46.110 POST /MotoList.aspx lng=th&make=452&model=31&ccmfrom=120&priceto=100000 80 - 180.183.49.146 Moz
2017-03-10 00:00:05 88.198.46.110 GET /MotoList.aspx lng=th&make=452&model=31&ccmfrom=120&priceto=100000&sort=model 80 - 180.183.
2017-03-10 00:00:05 88.198.46.110 GET /ListCounter.ashx ids=13925&rnd=839216 80 - 180.183.49.146 Mozilla/5.0+(Windows+NT+10.0;+Wi
2017-03-10 00:00:07 88.198.46.110 GET /MotoList.aspx lng=th&sort=latest&top=50&page=5 80 - 182.232.231.37 Mozilla/5.0+(iPhone;+CP
2017-03-10 00:00:07 88.198.46.110 GET /MotoList.aspx lng=th&s=%2fHONDA&top=250&make=130&page=5 80 - 58.8.154.133 Mozilla/5.0+(Lin
2017-03-10 00:00:07 88.198.46.110 GET /ListCounter.ashx ids=15033,15029,15027,15024,15023,15018,15017,15016,15015,15014&rnd=48797
2017-03-10 00:00:07 88.198.46.110 GET /ListCounter.ashx ids=15141,15140,15139,15138,15137,15136,15135,15134,15133,15132&rnd=68646
2017-03-10 00:00:07 88.198.46.110 GET /Photos/Small/0058/883.jpg - 80 - 182.232.231.37 Mozilla/5.0+(iPhone;+CPU+iPhone+OS+10_2_1_
2017-03-10 00:00:08 88.198.46.110 GET /ListCounter.ashx ids=15173,15129,15126,15118,15093,15069,15055,15052,15045,15004&rnd=43233
2017-03-10 00:00:08 88.198.46.110 GET /Photos/Small/0058/567.jpg - 80 - 58.8.154.133 Mozilla/5.0+(Linux;+Android+5.0.2;+vivo+Y35+
2017-03-10 00:00:08 88.198.46.110 GET /Photos/Small/0058/979.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:08 88.198.46.110 GET /Photos/Small/0058/871.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
2017-03-10 00:00:08 88.198.46.110 GET /Photos/Small/0058/882.jpg - 80 - 182.232.231.37 Mozilla/5.0+(iPhone;+CPU+iPhone+OS+10_2_1_
2017-03-10 00:00:08 88.198.46.110 GET /Photos/Small/0058/869.jpg - 80 - 49.230.224.223 Mozilla/5.0+(Linux;+Android+5.1;+iris+550+
  
```

A "Find" dialog box is open over the text area. The "Search for:" field contains the text "/Photos/Small". The "Replace by:" field is empty. The "Start search:" options are "Beginning" (unselected) and "Current Position" (selected). The "Case Sensitive" checkbox is also unselected. The dialog box has buttons for "Search", "Replace", "Replace all", and "Close".

Kainet Editor can be downloaded from our web page. It is free for private users, self-employed persons and small companies.

## 11.2 Kainet Aspi

Kainet has developed a complete solution for blocking spiders. It is a http module which can be integrated into ASP.NET web sites.



There are plans to implement the solution as a reverse proxy, so it can be used platform-independently.

## 11.3 CFLogFileConverter

This tool can be used if the logs are provided by Cloudflare.

- ◆ A Cloudflare log push will create a folder for every day, containing several thousand compressed chunks containing log files
- ◆ The converter reads these chunks directly and then creates a traditional log file
- ◆ During this process the log entries are sorted, as the original sorting is rather bad and can jump back and forward by several minutes

The software can be modified and optimized for various needs.

## 11.4 LogFileDownloader

This tool is used in load balanced environments and covers the following process:

- ◆ Download of zip compressed logfiles from different servers with FTP
- ◆ Decompression
- ◆ Merging them chronologically
- ◆ Produces one single output file

Also this software can be modified and optimized.